Material Identities in Corpus-Based Algorithmic Improvisation

Riccardo Ancona

Master Thesis

Institute of Sonology, Koninklijk Conservatorium Den Haag

2022-2023

Page intentionally left blank

Table of contents

Acknowledgements

I would like to thank Ji Youn Kang for her mentoring: her invaluable critical guidance helped me to question essential aspects and endure the research process. I also need to express my gratitude to all the other professors of the Institute of Sonology for their profound insights, especially to Peter Pabon and Raviv Ganchrow for instigating paradigm shifts in my thinking. Thanks to Francesco Corvi, Giulia Francavilla, Elif Gülin Soğuksu, Andrejs Poikāns, and Hugo Ariëns for sharing the stage or the studio with me. Thanks to Giovanni Onorato for his inspiring ideas. Thanks to my family for supporting me in these two years. Lastly, thanks to Chiara Buoni for always being present and encouraging.

Section A: A Theory of Perceived Materiality

Chapter 1: A Phenomenological Premise

A phenomenology is a description of how a reality manifests itself to consciousness. It is not our interest here to discuss the objectivity of such a reality: as Carl Gustav Jung (2014) observed, if it were a psychic reality, it would still be a reality for the subject experiencing it. Phenomenology starts from the experiential, intrapsychic dimension of the subject and 'brackets' the world by suspending any judgement on its ontological consistency. Since the aim of our phenomenological reduction is to understand auditory phenomena, the notion of suspension of judgement – or *epoché*, as Edmund Husserl (2012) called it – is not to be understood as a universal removal of all prior knowledge of the world, but in terms of a 'local epoché' aimed at questioning subjective experience, without presupposing that the contents of such experience physically exist.

Thus, if we try to put the world in brackets, close our eyes and concentrate on what we hear, we will find ourselves in a situation similar to that desired by Pierre Schaeffer (2019). Several objections have been made to Schaeffer's reading of phenomenological reduction (Solomos 2000), from which he derives the notion of reduced listening, according to which it is possible to practice a form of listening where all causes of sound have been neutralised in order to concentrate on the 'sound in itself.' My standpoint is that there is no 'sound in itself'; as François Bonnet (2016, 112) states in his critique of reduced listening, 'it is impossible to disentangle a pure sensible from a pure meaningfulness. […] an objectified sensible is already meaningful.' We can play at deceiving ourselves by trying to conceal the causal links, but even in the most detached, platonic approach to listening there remains an unwavering residue, an intrinsic consistency of sound that binds it to a material, tangible reality. Under no circumstances can the acoustic context around us disappear; it always remains distinct. The sense of hearing constantly informs us of the space, movement, events, forces, objects, and actions that populate the world in which we are immersed.

If I close my eyes now, trying to limit my attention to hearing only, I feel that I am here. I hear the static sound of ventilation systems from the next room, the hissing fans of the computer on which I am writing right in front of me, and the mechanical tickling of the keyboard, inharmonic trajectories made by motor vehicles outside the window, the distinctive call of a dove (precisely, a Streptopelia turtur), the intense friction of a drill piercing a solid surface. I may be distorting the reality of things, perhaps the bird I hear singing is not a Streptopelia, but I wish it were, because it is my favourite bird; perhaps there is no bird, yet it appears as immediate data to my consciousness. As to be discussed in the following chapters, these immediate data to consciousness go beyond the realm of phenomena; they are already interpretations that we inevitably perform. It is naive to think that by bracketing the world it is possible to unlearn what has been structured in a lifetime of listening. In every instance of adult acoustic perception there is a pre-perception that we cannot deprive ourselves of (Chion 2016). The spontaneous inclination of the subject to delineate an acoustic context cannot remain dormant for long. As a subject, I feel the need to know where I am, what is happening, what is going on around me, what forces govern the surrounding space, what comes before and what after, comparing the present with my memory.

In the context of this thesis, the spontaneous question of main concern is 'what?', the *ephemerals of the what* (Ganchrow 2021) that continually frame our subjective experiences. It may be possible through phenomenological reduction to set aside the ontological consistency of the world, but it is not possible to detach the sound experience from the need to formulate an acoustic context around the subject and designate a field of forces and entities – whether those entities are revealed to physically exist or not. What is fascinating is to understand how these entities frame themselves in our intrapsychic dimension, since every acoustic event contains a "what?", a question mark to which the subject feels the urgency to formulate an answer. Our very faculty of perception confronts us with a constant enigma that we cannot help but investigate; a gift, but also a problem for us to solve, to put it in Wittgenstein's terms (Kosuth 1990).

Yet it is not merely a process of rationalisation. The subject is not primarily looking for logic, but for an orientation, for coordinates within which to place itself. Note that in the case of the acoustic experience, the term 'coordinates' should not be collapsed with the visual notion of geometric coordinates: it is not a matter of abstractions, but of contingency, and situatedness. It is not absurd to consider the possibility that situatedness may be the most accurate definition of consciousness. Unsurprisingly, the *Phenomenology of Perception* by Merleau-Ponty (1945) pinnacles on an inter-corporeal relationality between the subject and the world – the situatedness of the body may be understood not just as a compass, but as the intrinsic core of this transcendental relationship. I argue that that there is no coherent notion of consciousness outside of this situatedness and the inhabiting of the body and its surroundings. Such a position consequently implies a rejection of both aprioristic idealism and Schaefferian attempts to separate sound from its context; the experience of the phenomenological subject cannot deprive itself of its context, since it is constitutive of the subject itself. Moreover, such a perspective also eliminates the risk of falling into a solipsistic interpretation of reality: the constitutive transcendental character of the situatedness of the subject immediately conjures an element of coexistence with other subjects, a co-habitation, a 'mutual relationship of existence' in which we discover ourselves, as enunciated by philosopher Watsuji Tetsuro (1961, 5). It is because of this original interdependence that it could be fruitful to inspect the data of perception as possible structural elements of musical expression.

The act of listening informs us of the concreteness of our situatedness. Even in conditions specifically designed to neutralise or nullify the acoustic context, listening attempts to orient itself. In the darkness of the anechoic chamber of the Technische Universiteit of Delft, I did not have the Cagean epiphany about the non-existence of silence: my body was too intent on listening to itself in a space without coordinates, causing a visceral, macroscopically hallucinatory sensation of its own consistency. In an anechoic chamber, even a sine wave manifests itself to us in all its corporeality. Even if we bracket the existence of the world and induce disorientation in ourselves, the world insists on surrounding us from the inside and the outside in its tangible concreteness. The experience of this concreteness, not its actual physical existence, I call *perceived materiality*. From this phenomenological intuition, a research trajectory is established that starts from the experience of the subject and through new theories and methodologies arrives at a particular musical approach. The next chapter will be concerned with defining the notion of perceived materiality, comparing it to the preexisting vocabulary of electroacoustic music. After this notion will be established, in Chapter 3 the phenomenological intuition will be scrutinised by contemporary paradigms of neurobiology of hearing and psychoacoustics, formalising a cognitive model of the learning process of *material identities*. An integrated theory of perceived materiality will be formulated in Chapter 4 in relation to a non-reductionist epistemological model and semiotic notions. Section B of the thesis, beginning in Chapter 5, will be concerned with defining a technical and practical methodology through which material identities can be represented in the computational domain, and understanding how perceived materiality can be conceived as a fundamental constructive aspect of real-time musical processes. The entire practical section will focus on the use of musical strategies of materiality in algorithmic musical improvisation, with a particular interest in collaborative performance. Chapter 5 will motivate the methodology employed in detail, while Chapters 6, 7, and 8 will describe the technical implementations of corpus-based strategies of perceived materiality. The development of techniques, algorithms, and models will form a performative environment whose characteristics will be commented in Chapter 9. Hopefully, the research path undertaken in this thesis will instigate new reflections on sound, instil criticism, and demonstrate how it is possible to engage a specific attribute of perception and investigate it with the aim of deriving a musical praxis from it.

Chapter 2: Terminology

The concept of perceived materiality needs a definition. It is not a pre-existing concept in the literature of electroacoustic music, although there is a constellation of notions having a tangent meaning. None of these notions is totally equivalent, nor does it seem to be able to exhaust the horizon of creative possibilities deriving from an inspection of this concept. Why is there an urgent need to create a new concept? Isn't the lexicon of sound perception already bewildering enough? A logician might argue that 'entities might not be multiplied beyond necessity' (Crombie 1959, 30), but what is a necessity in the domain of artistic research? If, according to Deleuze and Guattari (1996), the role of philosophy is to invent new concepts, such a vocation is to be found all the more in artistic literature. In this circumstance, to invent concepts is not merely nominalistic, but heuristic: it is not a matter of naming things, but rather of discovering new perspectives within which to interpret aesthetic experience. New methods and sensibilities may emerge from such perspectives – reversing Stockhausen's famous statement (1972), 'new methods change the experience', sometimes it is a new understanding of the experience that generates new methods. Therefore the subjective urgency to arrive at an unprecedented interpretation of the sensible is a sufficient reason to 'multiply the entities.' It will be the ability of the new concept to open up fertile ground for artistic production that will decree its legitimacy a posteriori. Take Schaeffer's (2019) notion of sound object: it is utterly fallacious both ontologically and phenomenologically, yet it has been an essential conceptual building block in the development of musique concrète. Its validity derives more from its effect on artistic thought than from any actual argumentative rigour. This is certainly not to say that the notion of perceived materiality is going to be fallacious at the outset; on the contrary, its scientific and epistemological consistency will be discussed extensively in chapters 3 and 4. For the moment, it is important to dwell on it as a phenomenological intuition, the nature of which needs to be circumscribed and specified.

When one attempts to delimit a sensible intuition such as perceived materiality, a palpable difficulty arises in defining it. The data of experience are not 'things' to which one can simply give a name in order to establish a biunivocal correspondence between object and word. At the same time, neither are they abstractions like mathematical concepts, for which one can define a semantic convention and elaborate a logical proof. Rather, they are qualitative experiences that by their very nature elude the concept of linguistic conventionality. Yet each of us individually experiences qualities first-hand in a dissimilar but certainly analogous way to what another human individual might do. The difference between the experience of one human phenomenological subject and another is probably surmountable – or at least, for the time being, it will be asserted that it is so, for the sake of a communicative optimism without which one would not even write theses. It is therefore necessary to continue striving to perfect a language made up of more or less stable references to the domain of subjective experience, so that a collective discourse could emerge from it.

Paradoxically, the closer what we want to define comes to the banal immediacy of experience, the more complex it is to talk about it. A tangible effect of this nominalistic difficulty can be observed in the terminological plurality that spontaneously develops when we want to refer to a subset of sensible experience. In the terms of the discourse of electroacoustic music, an overwhelming multiplicity of terms can be found, all pertaining to the semantic sphere of perceived materiality: *materiality* (O'Callaghan 2015), *physicality* (Wishart 1993), *organicity* (Dufort 2008), *concreteness, material potential* and *material individuation* (Döbereiner 2019), *tactility factor* (Anderson 2011), *sonic surface* (Harrison 2000), *concrete materiality* (Stavropoulos 2019). These terms are generally used – taking their meaning for granted – often more for the need to evoke a vague domain of subjective musical experience, than to refer to a defined concept. Even Michel Chion, who has made remarkable efforts to disambiguate and expand the lexicon of Schaefferian tradition, limits himself to a concise, inexhaustive definition when introducing his concept of *materializing sound indices:*

> [the materializing sound indices] denote aspects of a given sound that make felt more or less accurately the material nature of its source and the concrete history of its emission: its solid, gaseous, or liquid nature; its material consistency; the accidental features that arise during its

> unfolding; and so forth. The materializing sound indices of a sound can be greater or fewer in number and, in limited cases, a sound can have none (Chion 2016, 267).

As far as I have been able to ascertain, Chion has never deepened the concept. Although it may be a good starting point for reflecting on the subject of the materiality of sound, it is necessary to emphasise some problematic aspects of this definition. The indices of sound materiality would be first of all *indices*, therefore, by definition, something to which a quantity or a dimensionality can be attributed to; consequently they are quantitative properties that refer to qualities. Nevertheless Chion's definition does not specify in which physical quantities of the acoustic waves the dimensionality of these indices would reside, but only that there is a causal relationship between the intrinsic characteristics of sound and certain perceived qualities. The correlation between physical phenomena and qualitative perception of the subject is not explained, as the transcendental character of this relationship is not made explicit. While for Chion it is the sounds themselves that 'make feel' their own intrinsic material characteristics, the concept of perceived materiality emphasises instead the active role of the phenomenological subject.

Perception is not a unidirectional, passive act in which the subject deterministically absorbs the unambiguous characteristics of sound; on the contrary, it is a two-way, complementary process in which the phenomenon is actively interpreted by assigning attributes to it. The nature of these attributes and their actual relation to phenomena is by no means based on a deterministic principle, but is constantly shaped by the subject's prior experience, the hermeneutic mode with which they listen, their desires, and the surrounding acoustic context, as Chion also acknowledges in his *acoulogical* theory. These shaping factors cannot be removed from the experiential process, so it is reductive to attribute indices to acoustic phenomena as if they were unambigously revealing themselves to the listeners, all the more so if these indices would appear to have no traceable quantitative evidence in the acoustic phenomenon itself. Of course there must be an inherent *something* in the acoustic wave that informs us of its material nature, but this something would not be sufficient to explain the phenomenological experience of perceived material properties, because beyond the physical event lie a myriad of neurobiological and hermeneutical processes that influence the subject's interpretation.

There is, in short, a difference between physicality – that is, the physical being of the acoustic wave – and perceived materiality, which is what the subject creatively interprets of the material features of the phenomenon. To confuse or deny that there is a substantial difference between the two notions, would mean to endorse a physicalist interpretation of reality that has so far been unable to fullfil its scientistic promise of explaining human experience solely through the objective representations of science. For the purpose of artistic research, a purely physicalist view of perception is sterile.

The definition of perceived materiality that is about to be enunciated, in its perceptual attribute, gives prominence to the listeners in their subjective and irreducible experience. For this reason, in no case should the concept of perceived materiality be confused with the generic use of the term 'materiality' to designate the physical domain of things. In this thesis, there will not be any reference to things or objects – this is not a treatise on ontology, nor it is intended to be – but always to the phenomenological experience of such supposed objects, since it is in the domain of experience that art manifests itself. This aspect is particularly evident in electroacoustic sound art, where there is often no tangible artefact to appreciate. It is the ability of sound to communicate *intangible tangibilities* (or rather, *tangible intangibilities)* that motivates an inspection of phenomena not from a merely acoustic, but auditory and sonorous perspective.

The aforementioned phenomenological aspects are captured with great awareness by Erik Nystrom (2013 and 2017) in his description of *textural materiality*. Nystrom's notion is inextricably linked to the concept of *texture* in the spectromorphological tradition, an abstraction to which there is no need to bind ourselves here. However, it is pertinent to denote that in Nystrom's understanding, textural materiality consists of the set of characteristics perceived in a texture that refer to physical properties of objects or entities, with particular regard to their surface:

> The *materiality* of spatial texture may be seen as the window into our experiential bank of perceptual interactions with physical media, where

> we draw from multiple sensory modalities in the process of diagnosing the characteristic of sonic "stuff" (Nystrom 2013, 23).

This definition encompasses a multitude of aspects that will be explored in the following chapters, from the experiential nature of sound recognition mechanisms, to the cross-modal aspects of perception. For the moment, it is relevant to grasp the relationship between materiality, perception, and the diagnosis of 'sonic stuff' – the *ephemerals of the what* mentioned in the previous chapter. Nystrom seems to suggest that the perception of the materiality of sound emerges from the subject's attempt to identify a source, or at least to circumscribe its physical properties. In this process, a series of qualities that I call *perceived material features* manifest themselves to the listener. Some examples of perceived material features are hardness, roughness, elasticity, surface tension, state of aggregation, as well as some more defined *material identities*, such as whether a sound recalls glass, wood, or metal. Although these qualities are attributed to the sound itself, it is the listener who performs a subjective process of attribution based on their prior experience of the physical world and on the current acoustic context.

By generalising Nystrom's insights outside the spectromorphological paradigm, it is possible to arrive at a first, provisional definition of perceived materiality: it is the cognitive faculty of attributing a set of perceived material features or a material identity to a sound event by the listener on the basis of neurobiological, experiential, emotional, semantic, and contextual processes, the evolutionary purpose of which is generally to recognise the identity of the surrounding acoustic signals.

What kind of considerations can be made about the nature of perceived materiality will be the subject of the first section of the thesis; on the basis of these considerations and of a defined methodological project, the concluding section will consist of a series of improvisational strategies that make use of the materiality of sound as a primary constructive principle. Before that, however, it is necessary to reflect on how perceived materiality emerges in the listening process.

Chapter 3: Neurobiology and Psychoacoustics

Having clarified the concept of perceived materiality of sound and delimited its boundaries in the transcendental domain of experience, it remains to be understood how this faculty develops in the human individual and how it functions. If our interest is in the phenomenological dimension in which material identities manifest themselves, then it is relevant to inquire what physiological, perceptual, and cognitive mechanisms and processes are involved in such manifestations; the question then arises as to whether, in addition to the immediate phenomenological understanding of experience, it is possible to resort to the scientific disciplines of neurobiology and psychoacoustics, in order to provide a foundation and explanation for what appears to us to be self-evident, yet elusive.

Perception studies provide an interpretative framework of the process of identification and signification of sound identities that is both indispensable and incomplete: indispensable because theoretical models of perception, supported by empirical research, describe the concrete dynamics by which identities take form and shape experience; incomplete because these sciences are still in their infancy and suffer from the limitations of paradigms that are still too narrow and methodologies that are far too inaccurate. In the sciences of perception, the field of study – the human mind in relation to its points of access to the world, the sensory extensions – is more than a mere field, it is a galaxy of dynamic relationships whose complexity far exceeds the current capabilities of the analysis tools and interpretative models employed. Nevertheless, current models of the auditory system and related cognitive processes already provide a useful orientation for understanding the perceived materiality of sound.

First, it is necessary to question whether the faculty of recognising material identities or even some attributes of perceived materiality is an innate faculty. Studies on the development of the human auditory system suggest that there is nothing, or nearly anything innate about it, since at the time of birth both the physiological apparatus and its organisation by the nervous system are not yet fully developed. In the first six months of life, the cochlea, the auditory nerves, and the neural connections in the brain responsible for the transmission of auditory information are still developing. From six months until the end of pre-school age, the neural circuits continue to mature and the most radical changes are observed in the elementary faculties of auditory signal processing, such as the segregation or grouping of sources and the ability to discern differences in time, intensity, and position of sounds. It has already been known for decades how 'the auditory system of adult mammals is both structurally and functionally influenced by auditory experience during early postnatal development' (Aislin 1981). Even once the physiology of the auditory apparatus has reached a form comparable to that of adults, this is not enough to make children's hearing equal to that of grown-ups. Werner observed that

> it is possible that the codes for many acoustic properties are available to the child, but the child has not yet learned to use those code efficiently', and that 'it is even possible that precision of auditory coding at the brain stem level improves as a result of top-down influences as the child discovers the utility of certain sorts of acoustic information (Werner 2012, 7).

It is likely that these improvements can also be attributed to a gradual refinement in the processing of information by the central nervous system (Buss et al. 2012).

The process of refining auditory and cognitive capacities can be understood as a form of perceptual attunement, in which the subject constructs its own orientation in the surrounding environment and learns to decode it with specific strategies; this attunement also takes place with regard to language (Panneton and Newman 2012). It is indubitable that the structuring of the linguistic domain imprints a profound mark on most of the faculties of listening, including the identification of the material properties of sound, but even beyond the linguistic imprinting on signification, the subject's empirical experience of the world has an undoubtful influence on the auditory faculties. In order to accomplish what psychoacousticians call *sound source recognition*, it is essential that the source to be recognised is an entity already present in the listener's database of experiences. It requires the listener to have had prior exposure to such a source, to have observed its intrinsic characteristics several times in the past with multiple variations – to be able, therefore, to trace the present experience back to a longer chain of recurrences in time, of which the current manifestation is a single instance of a broader phenomenon belonging to the alleged laws and norms deduced from the surrounding world. According to Albert Bregman (1990), auditory segregation and source recognition is based on two concurrent and complementary processes, the 'primitive processes', also called 'bottom-up processes', which consist of the decomposition and grouping of information extracted from the peripheral auditory apparatus on the basis of the acoustic properties of the signal, and the 'schema-based approaches', the top-down strategies that occur in the auditory cortex and are related to the listener's preexisting memory. As Yost explained,

> 'physical attributes such as temporal onset differences, harmonicity, and common modulation may be used by central nervous system circuits to segregate the neural code to account for sound source perception. At the same time, information gleaned from one's experience in processing sound sources can guide these neural circuits in processing the peripheral neural code' (Yost 2008, 4).

With regard specifically to the identification of the materiality of sound, the studies conducted so far are rather limited and have not produced yet any particularly surprising findings, although they do seem to suggest that no specific bottom-up element is sufficient in itself for this task. As stated in the previous chapters, no measurable acoustic feature has a direct and unambiguous correspondence with perceived materiality, precisely because there is an interplay between the physical signal and the neural structures which organise, compare, and categorise the signal according to memory. In his inquiry of material identification from sound, Lufti (2008) found that 'measures of performance accuracy have so far demonstrated our capacity for identification, but they have not permitted strong conclusions regarding the basis for identification' – confirming only what was already evident from a phenomenological perspective. The limitation of the research conducted so far on the subject is that it is usually run in the laboratory, on a small number of voluntary subjects, using substantially decontextualised material identities, largely nullifying the intrinsically transcendental character of perceived materiality. I am convinced that it will only be possible to better understand the functioning of material identification when this process will be studied in different contexts and in relation to a plurality of concomitant stimuli; meaning does not emerge from a single factor, but from a complex of coexisting elements.

Bregman's theory and related empirical research supports the idea that schemes exist at the level of the auditory cortex and perhaps in other memory areas that our brain employs, to quickly recognise previously formed identities. The precise physical conformation of such schemes is still unknown, as it is complex to disentangle the structures of high-level cognitive functions in an area as dense and articulated as the cortex. The establishment of schemes is necessarily subjective, resulting from a plurality of lived experiences, as well as undoubtedly influenced by the interpretative framework provided by the surrounding community. It is worth noting that the schemes theorised by source recognition – of which what I call 'material identities' can be considered a particular subcategory – are based on a plurality of prototypes that the individual encounters in the course of their existence, in particular of their youth. These prototypes may be rather contradictory to each other; indeed they would be partially incompatible if their acoustic characteristics were analysed and compared, yet the subject groups them together in a specific scheme, as this greatly facilitates and speeds up the onerous cognitive effort involved in listening. To some extent, theories of perception seem to confirm Chion's (2016) assumption that our ordinary perception mostly takes the form of pre-perception, as the cognitive load of listening is lightened by previously formed identities and patterns, in a way that shapes and sometimes even distorts the physical signal. The material identities through which we filter the surrounding environment are based on a catalogue (or rather a *corpus*) of experiences that the subject constantly consults and often updates, even after the developmental years, albeit with less plasticity.

If we strive to remember how complex and demanding it was to establish in ourselves the basic categories of material reality, we cannot help but realise how concrete and embodied these experiences were. When infants are still at the beginning of their process of framing the world around them, they have a whole physical reality to motivate and structure. It is not uncommon to observe infants playfully dropping objects to observe their trajectory, learn their response to impact and, crucially, experience what that body sounds like. I have the impression that playing is the primary horizon within which one learns how material reality behaves and what sounds it produces. In playing, however, as in most human activities, there is more than just listening: playing also involves touching, looking, smelling, and tasting. Much of our experience of the materiality of sound depends on the tactile, visual, gustatory, olfactory experiences that complement it, because perception is essentially cross-modal, as cross-modal is our way to acquire information from the surrounding world. Hence, it is difficult to determine exactly how the signification of perceptual data takes place, as the idea of a unified framework between the various disciplines of perception is not yet the dominant paradigm. Philosopher Casey O'Callaghan, starting with the study of cross-modal illusions, has formulated a theory of perception which expresses the absolute need for a multisensorial approach. He argues that

> the coordinated use of multiple senses enhances and extends human perceptual capacities. […] It can enable us to perceive intermodal identity and novel intermodal instances of relational features (O'Callaghan 2019, 198).

It does not seem inconsiderate to believe that material identities are intermodal identities, since the interaction we have with matter is intermodal. In this sense, a study of the perceived materiality of sound without a study of the perceived materiality of the other senses will always be limited. Certainly when the various branches of the sciences of perception will be unified and will begin to be interested not only in reduced case studies on simplistic acoustic phenomena, such as sinusoids and pink noise, finally inquiring the listening subject in its situatedness and in relation to the complexity of the acoustic reality, then science will have much more to tell us about perceived materiality. Even in this primordial stage, the results of empirical research on sound source recognition nonetheless provide us with an orientation that will be essential for the development of this research.

Chapter 4: Epistemology and Semiotics

The interpretative models of perceived materiality proposed so far are not exhaustive. There are other perspectives from which material identities can be analysed to achieve different forms of understanding. Each speculative discipline has potentials and structural limitations, but once interconnected with the others, it contributes to form a multidimensional view of the subject under examination. Rather than relying on a unified theory of perceived materiality, it is more fruitful to attempt to observe it from several divergent points of view; after all, what is pertinent is not its ontological unity, but its properties and the creative possibilities derived from them. Hitherto, attention has been paid to the phenomenological perspective and the empirical-physiological theories. These two methodologies of analysis respectively provide one with the tools to identify perceived materiality as a subjective phenomenon and the other with a structured framework to trace its biological functioning in the individual. There remain numerous further possible points of access that could be undertaken to gain a deeper understanding; in this thesis, I will engage two more speculative approaches in order to clarify certain inherent properties of perceived materiality. Nevertheless, the aim of this analysis is not completeness, but the heuristic potential of a multidisciplinary perspective. The perceived materiality of sound will therefore be briefly investigated from an epistemological and semiotic perspective.

Epistemologists are concerned with asking what kind of knowledge is possible, rationally verifying the degree of truth embedded in experiences and statements. Epistemology thus aims to delimit knowledge and observe its boundaries. The gnoseological properties of the perceived materiality of sound can be studied in very different ways depending on which epistemological orientation is adopted. The principles of non-reductive physicalism will be followed. With regard to the perception of phenomena, this perspective holds that there is only one type of ontological substance, namely physical matter, denying the existence of further mental or ideal metaphysical constituents; at the same time, non-reductive physicalism claims that scientific predicates are not sufficient to comprehensively signify the totality of experience, since although psychic experiences have a purely physical basis as well, there is no unambiguous correspondence between a given segment of experience and a determined neurobiological predicate (Stoljar 2022). In this sense, the non-reductionist perspective seems to favour emergentist hypotheses regarding high level cognitive activities such as the recognition of material identities. An emergentist perspective might even entail the skeptical thesis arguing that the factors causing such cognitive skills are so untangled that they will never be fully explained by means of the current scientific paradigm, as the awareness of reality involved in material identification is an emergent property.

In the field of epistemology, perceptual experiences are often considered *qualia*, a term used to refer to qualitative subjective sensations. The concept of qualia is open to numerous interpretations – see (Tye 2021) for a more thorough examination of the various currents of thought. Qualia are often labeled as epiphenomena, in the sense that they are byproducts of perceptual phenomena. The perceived material properties of sound fall within the realm of experiences commonly identified as belonging to qualia. Some philosophers, starting from Peirce (Lynn 1985) who first introduced this concept, argue that qualia are non-representational, since they reflect a feeling that precedes or is independent of a symbolic signification; for Peirce, perceived material properties are qualia, but not material identities, since identities are already structures holding a symbolic mark. As the neurobiological theories described in the previous chapter indicate, it is very difficult to separate the structures of signification from those of perception, ascribing pure feeling to a moment prior to that of the attribution of meaning; there is in fact a constant, mutual participation of sensory and symbolic levels throughout the field of human experience. Whether these two levels are simultaneous, subsequent, shared, split, or univocal is not the crucial point: what is relevant here is that, beyond the symbolic representation that will be analysed later by the means of semiotics, there are subjective qualitative epiphenomena coinciding to the particular perceived material properties of sound. Qualia are immediate to consciousness, intrinsic, and private, since they concern the singularity of the subject experiencing them. This means that although there is a common ground between subjects – the physical world and its materiality – qualia always retain a degree of incomparability. From an epistemological point of view, there is no way of establishing exactly what relationship exists between your qualia and those of a subject who is sharing the same surrounding reality with you. Given their private, qualitative, inherently subjective nature, qualia are incommensurable, which means they cannot be represented in quantitative terms. It is possible to measure the amplitude of the individual spectral components of a sound, but this measurement does not relate unambiguously to the qualia that the sound produces in the listener – this epiphenomenological byproduct is by definition incomputable, because although it consists of a specific conformation of the brain matter and its neuroelectric equilibria, it is not possible to derive numerical factors that correspond unambiguously to this experience. The incommensurability of qualia implies a structural problem in translating this theory of perceived materiality into the computational domain of algorithmic music. This epistemological limitation is the starting point from which the questions and pragmatic objectives of this thesis move. Before dealing with them, it is important not to overlook the semiotic interpretation of the subject under investigation.

Whereas so far this discussion has focused on perceived materiality as an essentially intra-subjective event, it is now critical to examine the inter-subjective and symbolic aspect of material identities. If the semiotic level is neglected, there is a risk of interpreting perceived materiality as an entirely private fact; on the contrary, communication between subjects shapes the schemes of representation of identities. A semiotic perspective on material identities entails studying them as semantic units, as signs. More accurately, since there are no signs disconnected from a broader system of sense relations, it is more appropriate to refer, starting with Peirce (1991), to *sign relations* rather than to signs. From an inter-subjective point of view, a material identity is a sign relation that encloses a certain class of perceived material properties according to a convention established by the cultural context. If in the section on neurobiology the focus was on embodied subjective experience as a constitutive element of the schemes of signification of reality, semiotics completes the ontogenetic framework by showing how the community shapes the discursive and interpretative categories of individual subjects. When children play with a material, they learn its properties directly; at the same time, the interpretive model that is formed in each of them, which is also semantic, is influenced by the discourse that adults and other children produce about that material. Semiotics allows us to think of material identities no longer as totally private monadic entities, but as structures of perception whose characteristics are largely dependent on cultural discourse. A sign relation is thus a scheme that certainly has a subjective interpretation resulting from the accumulation of perceptual and discursive experiences, but it is also a cultural unit that is embedded into a broader code of sign relations (Eco 2016). This code should not be understood as a static structure – on the contrary, Umberto Eco argued that

> the mobility of semantic space makes codes change transiently and processually. But at the same time it imposes on the activity of sign production and […] interpretation itself the necessity of a continuous *extra-coding*, [in a way that the listener] is at the same time obliged both to challenge the existing codes and to advance interpretive hypotheses that work as a more comprehensive, tentative and prospective form of codification (Eco 1976, 129).

In light of a semiotic perspective, the identification of perceived materiality is the result of the always dynamic relationship between perception, preexisting semantic structures, and interpretation, the latter understood as an inherently creative activity.

It is relevant to ask how sign relations relate to each other. If the meaning of a given semantic unit were to be defined, as in a dictionary, other semantic units would have to be called in to make up the definition; each of those units, in turn, would require the use of other units, and so on *indefinitely*. Meaning, as explained by semiotics, is a network of significations interconnected according to a process of infinite semiosis (Atkin 2023).

Paradoxical as it may seem, our discursive horizon is a self-referential system, a network in which the new codifications that continually flourish are recursively articulated in reference to preexisting cultural codes. Therefore, the definition of a material identity as a cultural unit is part of a broader, infinitely recursive process of signification. To navigate the networks of meanings informs us more about the nature of symbolic systems than the nature of material identities; what can be explored of material identities from a semiotic point of view informs us of the discursive system of reference, but does not allow us to access deeper layers of the neurobiological schemes that produce the experience of such identities.

Both epistemological and semiotic perspectives, while offering meaningful points of view, provide us with an interpretative framework that appears problematic. What kind of representation can be conceived of perceived materiality if it manifests itself as a para-subjective dimension, scarcely comparable between subjects, indefinable without recursion, inherently incommensurable? What kind of knowledge is possible of such an elusive notion? Is a systematisation of material identities possible? The hypothesis of a taxonomy of perceived materiality will be examined. Biologists employ taxonomy as an ordering tool for the multiplicity of the organic world, structuring a classification according to morphological and genetic criteria. Taxonomy is a way of organising phenomena that constructs networks of relationships taking into account the degree of kinship, similarity, and derivation between the various elements of the system. When faced with the plurality of material identities, it might be logical to envision ordering them in taxonomic terms, but it is difficult to determine which ordering criteria can be employed. A typo-morphological approach could describe certain fundamental properties with which to distinguish and relate material identities. Dichotomous pairs could be taken, just as is done in biological taxonomy, establishing dual opposites: hard/soft, smooth/rough, concave/convex, light/heavy, solid/fluid, and so on. Any sound could be taken, listened to carefully in an attempt to establish its material identity by means of these dichotomous pairs, but structural difficulties immediately emerge. For instance, there are sounds that are neither smooth nor rough, and if a sound is fluid, there is no point in asking whether it is hard or soft, but rather how viscous it is; when one says that a sound is light rather than heavy, what metric is used to compare? If qualitative adjectives are used, it is only possible to refer to the shared semantic value of such adjectives, which, as discussed above, informs of cultural conventions, but does not tell much about the intrinsic perceptual characteristics of the phenomenon it attempts to describe. Instead, it would be necessary to define reference sounds against which the relative differences of each individual sound could be established, or it would be possible to compare each sound with another, attempting to form a scale or gradient of a given material property. Yet when attempting to systematise numerous sounds classified by these criteria, a series of increasingly problematic questions emerge. Ordering sounds by gradients of perceived materiality implies the idea that there is a quantitative factor inherent in the epiphenomena of materiality; however, it has already been explained why such a position is not acceptable on an epistemological level, as the qualia of perceived materiality are by definition incommensurable. If a hundred short sounds are taken and an attempt is made at ordering them according to a gradient, say from the hardest to the softest, one will soon experience how impossible it is to establish an ordered scale of relationships. Even assuming a gradient could be established, a new taxonomy performed in blind listening a few days later would produce a completely different order from hardest to softest, revealing how inconstant perception is when it comes to establishing quantitative or comparative estimates on properties as fuzzy and subtle as the perceived hardness of a sound. I tried to classify two hundred and fifty short sounds according to multiple gradients of perceived material properties by performing a linear regression algorithm, relating the recorded material properties to certain acoustic properties of the sounds in order to derive an equation; despite several attempts, the result of the regression is always indistinguishable from that of a pseudo-random numerical sequence. Material identities are based on prototypes, not vectors, and these prototypes are totally disordered in their inherent acoustic characteristics. It is therefore empirically evident that no constant gradient of perceived material properties can exist, and that an epistemologically based taxonomy is equally impossible. Of course, it is possible to construct completely arbitrary taxonomies as a means of freely ordering the chaos of reality – and this will be explored later – but such systematisations allow neither objectivising nor rendering computable what is systematised. The practical part of this thesis will mainly deal with the use of material identities in the computational domain – how can one make use of them if they seem imponderable to a machine? It will be a matter of establishing what kind of approximation is possible and what this approximation produces from a creative point of view, inventing a methodology that might take into account all the theoretical knowledge acquired so far on perceived materiality. Before focusing on the implications of this approximation, it will be necessary to linger temporarily on the practical context in which it comes into play.

Section B: Strategies for Corpus-Based Material Processes

Chapter 5: Methodology of Practical Research

The aim of my practical research is to employ the theory of perceived materiality described so far in the context of algorithmic improvisation. The question that drives the experimentation is how material identities can play a primary role in the design of algorithmic real-time musical procedures, behaviours, and processes. The goal is to find a set of logics through which it is possible to act with and upon perceived materiality in the domain of algorithmic improvisation, where computation and human thought form a single network of mutual feedback. The principles of the theory of material identities will be taken as an operational paradigm for the design of the performative environment. This undertaking is neither neutral, unambiguous, nor scientific. Rather, it reflects a series of deliberate and aware choices, some of which reflect personal taste and interest rather than objective necessity. The kernel of artistic research partly lies in the subjectivity of such choices. The techniques that will be described in the following chapters are based on a framework that represents a synthesis between the ideas described so far and a set of pragmatic orientations. It is therefore necessary to dwell temporarily on what informs my musical practice, as this will also have a substantial impact on the choices that are made in the development of techniques and strategies of perceived materiality.

Firstly, it is crucial to observe that I attempt to place my music-making at an intermediate and bridging point between an acousmatic-phenomenological approach and a structural-procedural approach. My major influences in the field of electroacoustic music tend to be roughly divided into two distinct strands: on the one hand, acousmatic music of spectromorphological derivation, and on the other hand, electronic music of structuralist or proceduralist origin. The acousmatic-phenomenological approach to which I refer flourished as a result of Schaefferian theories and later expanded with other paradigms stemming from those theories. The spectromorphology of sounds (Smalley 1997) and the degree of surrogacy of sound sources (Emmerson 1986) are strictly relevant compositional aspects. This framework of thinking places the relationship with the musical material at the core of the creative process, usually of acoustic origin and often containing semantic and metaphorical elements (Bayle 1989), and its transformation through sound processing. In this context, the fixed-media format is favoured, and the conception of musical time being adopted reflects and partly imitates the timing of film montage. Among the composers who employ this type of modality, Trevor Wishart and Manuella Blackburn have had a lot of influence on my music-making. Wishart (1993) has devoted much of his research to the concept of metamorphosis, in reference to an 'alchemy' of sound in electroacoustic music; he has also outlined a number of useful strategies for the management and proliferation of large masses of compositional material that I have been influenced by – although his approach involves offline sound processing. Manuella Blackburn (2019) has developed a microscopic approach to recording, processing, and editing of sounds with remarkable care for material identities. Her idea of employing a plurality of small sound tiles to holophonically compose larger structures is very prolific for me, even if I find its handcrafted modes of execution too linear and direct for my personal taste.

The structural-procedural approach derived from the most structuralist fringes of integral serialism, as well as from cybernetic theories, and then developed further according to post-structuralist paradigms. It involves focusing on the modalities and criteria by which generative processes are designed. The generating process and its musical potentials sometimes assume greater significance than the musical piece, as a fixed-media or performance is understood only as an individual instance within a statistical field of potential outputs. A structural-procedural composition is thus a possible configuration of a parametric space, in which stochastic and probabilistic techniques, as well as abstract or sometimes even arbitrary logical-mathematical procedures are employed. Since the focus is on processes, the musical materials must be as plastic as possible, which is why the use of synthetic sounds is generally preferred; samples tend to retain a previous history of their original emission, an history that remains inscribed in them against the proceduralist's intention, even after having been transformed by abstract procedures. Among the structuralists, Karlheinz Stockhausen certainly had a decisive influence on me, as he continually tried to define new criteria and new techniques for each work; in this way, the work had inscribed within it a matrix with which other potential works could be generated. From Stockhausen I inherit a fascination for the design of particular systems within which multiple outcomes can emerge – not one general system, but many systems, each with its own specific strategies and resulting musical possibilities. Among the more recent composers of structural-procedural music, I recognise as essential the work of Erik Nystrom, whom I have already mentioned in chapter two, for the way in which his practice finds a particular synthesis between algorithmic process design and post-spectromorphology.

Over the last few years, my intention has been to find an integration between the acousmatic-phenomenological approach that observes the nature of sounds as primary input, and the structural-procedural approach that articulates processes and events in a field of potential interactions. Then there are other poietic modes that have allowed me to broaden my point of view, problematise my influences, and define a praxis, such as live coding and post-acousmatic practices. Live coding, understood as the musical or intermedia practice in which the improviser alters blocks of code on-the-fly, has enabled me to think process-based music in the performative context. It has led me to privilege an approach in which composition consists of the design of the performative environment and its inherent possibilities (Magnusson 2014), while the formal development and the articulations of sounds depend on the choices made during improvisation. In live coding there is a substantial separation between performative gesture and resulting sound, which means that there is no coherent and unambiguous relationship between the two. The latency between gesture and musical outcome, although potentially problematic in a choreographic sense, implies a relative independence of the performer from the cognitive load deriving from the sensorimotor mechanisms frequently employed in other kinds of improvised music (Ancona 2020). The live coder harnesses a kairotic temporal dimension in which one waits for the most opportune moment to trigger change in the musical processes (Cocker 2018). This type of relationship is also possible in other forms of process-based musical improvisation and it is the performative mode I tend to prefer.

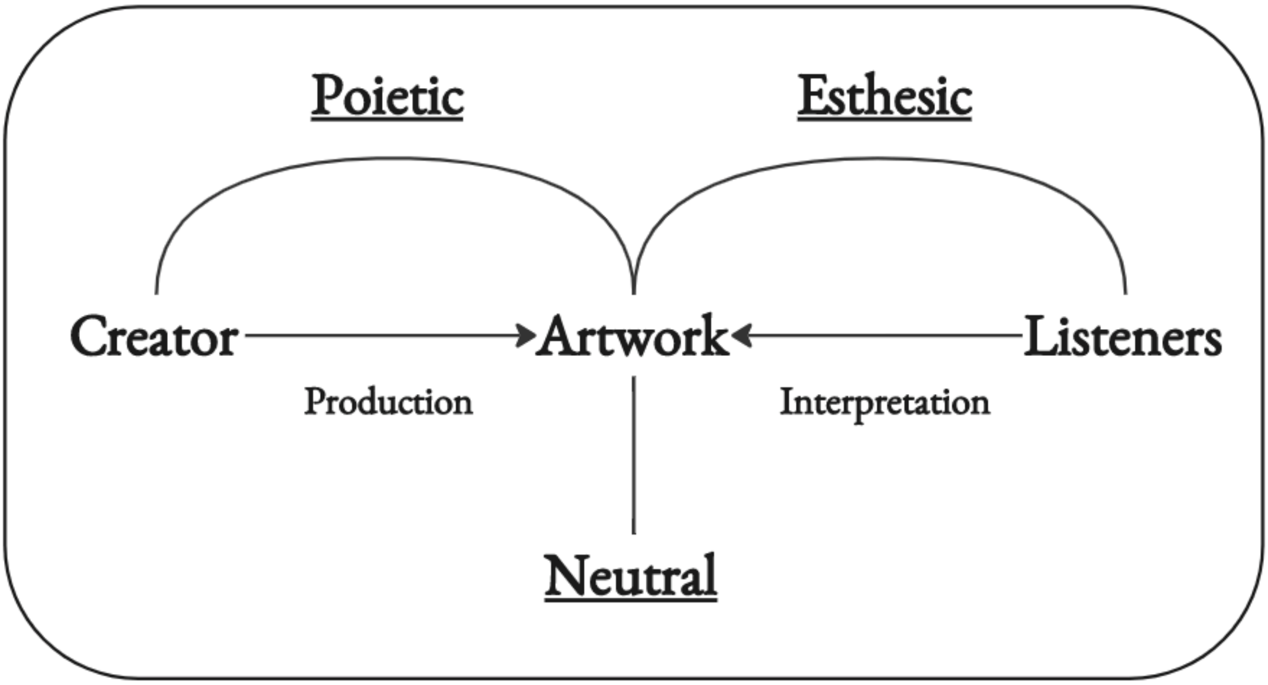
The term post-acousmatic, on the other hand, has begun to circulate in recent years in the English academic context, with a variety of contrasting meanings (Adkins et al. 2016); Onorato and I (2022) have proposed a definition of the term, according to which post-acousmatic practices are understood to be all those modes of music-making in which notions traditionally relevant in the acousmatic domain are transposed, instead, outside the uniquely aural condition. Whoever has inherited the criteria of listening, analysis, and composition peculiar to the acousmatic tradition, and then finds themselves applying these criteria in a broader sphere in terms of the perceptual layers involved – adopting different formats and modalities from the only-aural fixed-media – is practising music or intermedia art of post-acousmatic character. My approach to music-making can be defined as post-acousmatic, in the sense that from acousmatic music I inherit a certain focus on listening to sound sources and their characteristics from a phenomenological point of view, but at the same time I elaborate those forms with performance modes that diverge from the idea of a fixed-media composed as montage.

Borrowing from all the paradigms described so far, my personal synthesis involves:

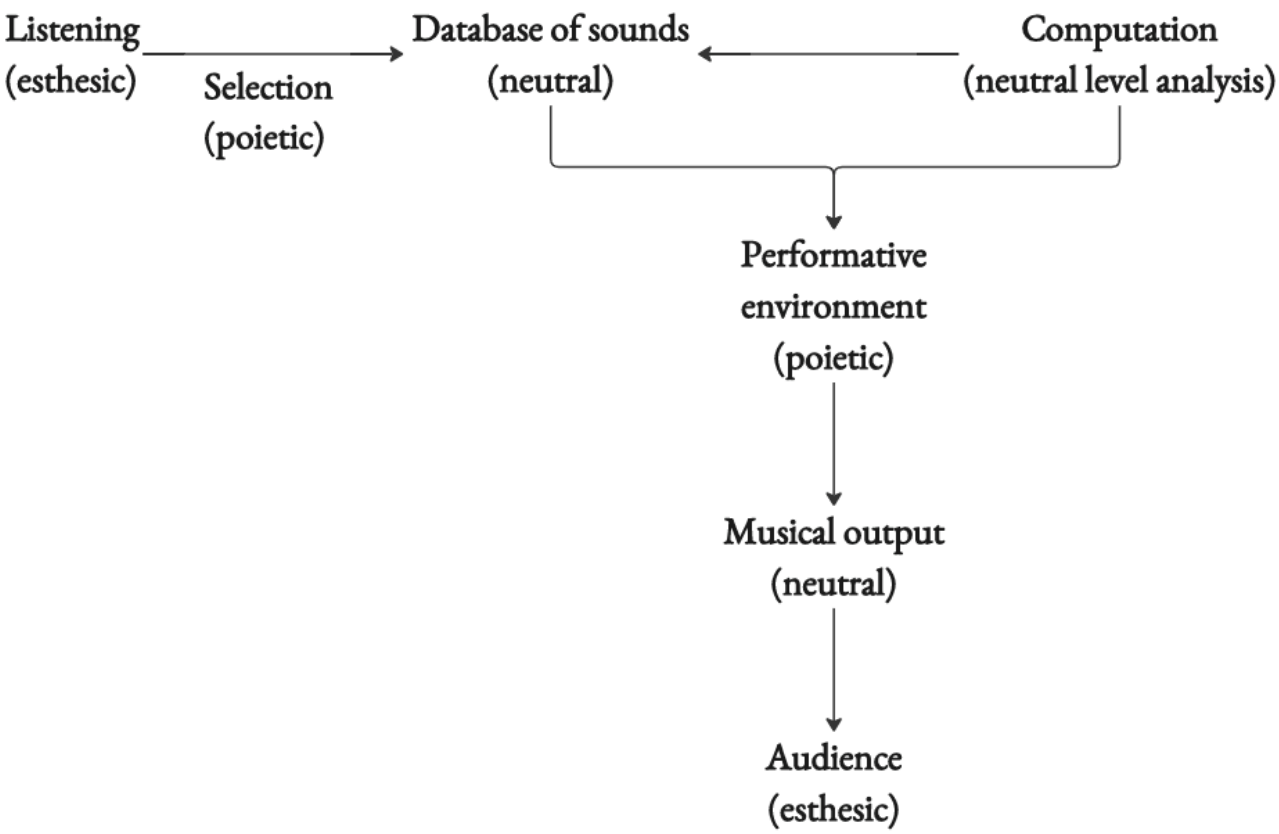
- A study of the perceptual characteristics of source materials inherited from acousmatic music;
- An interest in the semantic and metaphorical aspect of sound, also derived from the acousmatic tradition;
- The exploration of new algorithmic procedures for sound processing and articulation, borrowed from structural-procedural thinking;
- A focus on the real-time, performative, improvisational domain, acquired from live coding;

- The desire to design a performative environment understood as a system within which processes and networks of reciprocal relations can occur, an idea of post-structuralist inspiration.

For a deeper understanding of how my practical methodology is framed, it may be useful to employ a semiotic analysis model. Jean Molino (1990) devised a threefold scheme of symbolic forms, later employed by his student and musicologist Jean-Jacques Nattiez for musical analysis. Here, the tripartition will not be used for the analysis of a piece, but to illustrate the entire system of interchanges, causations, and relationships involved in my music. The figure shows the scheme of Nattiez's tripartition.



According to Nattiez (1990), there are three levels of analysis: the poietic level, the neutral level, and the esthesic level. The poietic level is the set of concrete ideas and strategies with which a work is produced by its or their author(s), the neutral level is the work in its concreteness, or rather the concrete trace that is intelligible to the senses, and the esthesic level is the set of interpretations produced by those who experience the artwork. The following scheme represents the process of my music-making as a whole in semiotic terms.

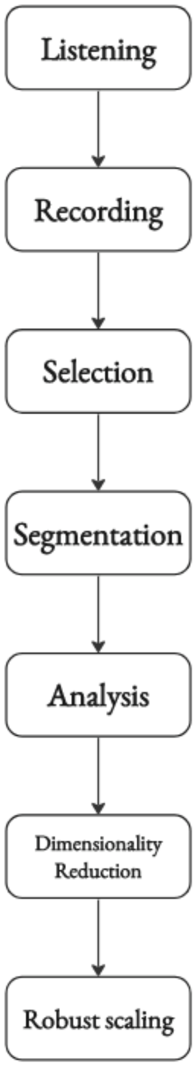


My creative methodology involves first of all an esthesic listening towards the surrounding acoustic reality; in this initial phase, listening is disinterested in the poietic potential of sounds, it is absorbed in experiencing their phenomenological characteristics. When I spontaneously establish a particularly meaningful relationship with a material identity, then my listening becomes poietic and I begin to imagine how that identity could be employed in my musical practice. The chosen material identity is then allegedly captured by recording it from as many perspectives as possible, and a set of recordings are selected as the starting musical material. The collection of sounds selected by the poietic action forms a new neutral level, a *corpus* of instances of a material identity. At this stage, my methodology involves a computational analysis of the corpus; the analysis of the machine could be understood as a computational esthesis, in the sense that – although the machine is probably not aware of its actions, at least in the same terms as humans are – the numerical analysis of a sound is a form of symbolic decoding of an acoustic event. Whilst the analysis of a machine is deterministic by design, there are a multitude of different potential types and configurations of analyses, which is why the human's choice to employ a given configuration is a poietic procedure that has a direct effect on the esthesis of the machine. Data obtained from the analysis, when correlated with the analysed sound corpus, form an extra-coding of the neutral level. I call such symbolic extra-coding of the corpus of a material identity a *material database*. From the material databases I construct, according to the criteria described in the following chapters, a series of processes, interactions, and musical strategies that together form my performative environment, the poietic space in which the musical output is produced. The music resulting from this poietic process consequently becomes a further neutral level, which in turn becomes the object of listeners' esthesis. Whether the listeners' esthesis involves an understanding of the poietic process or not is not a matter of concern. My fascination with material identities is esthesic at the beginning of the creative process and constitutes the phenomenological starting point within which poietic procedures are triggered. I am not always interested in inducing the listeners' esthesis to be focused on the perceived materiality of sound; in most cases I would say that I am not particularly worried by making the original material nature of the corpus intelligible to the audience, although there are circumstances in which I enjoy playing with the relative recognisability of acoustic sources and related metaphorical aspects. Superimposing the author's intention with the listeners' interpretation, having the presumption of conveying a message that clarifies poietic processes is a paternalistic perspective: for me, there is no message to convey – at least not in terms describable by other symbolic codes than the music itself – it is only a matter of configuring a system within which unheard sound relations might emerge, whose interpretation by the audience is totally open. Certainly, each poietic choice has the potential to produce a certain field of aesthetic results. Therefore, the author's expressive intention is inscribed in each choice. My curiosity in this context is not just aesthetic, but strongly speculative: I wonder which outcomes could derive from the assumption of a specific poietic framework. In this research, my interest is entirely focused on exploring what kind of potentials might emerge from the methodology I described. The application of these methodological procedures led to the development of a number of compositional, computational, and performative strategies that will be described in the following chapters.

Chapter 6: Material Databases

A material database is an ordered set of sound segments that conform to the same material identity. Each of the segments represents a singular instance of the material, whose acoustic characteristics were analysed and transcribed. The material database contains both the sound fragments and the data resulting from their analysis, so that a wide range of ordering and sound articulation operations can be performed. As in the human experience of materiality explained in the previous chapters, in order for material identities to take shape, there must be a multitude of associations between the physical phenomenon and a series of schematic representations of their characteristics. Likewise the material database has the task of translating this into the computational domain: the phenomenon to be observed is provided by buffers containing the amplitude information of sounds over time, while their synthetic representation consists of a description obtained from numerical analysis. In this way, the material database emulates the schemes of perception, but does so according to the logic of the machine, allowing for the elaboration of strategies for the articulation of material identities in the digital domain. The function of databases is therefore not to assiduously imitate human listening – which is in any case impossible – but to translate the cognitive process of perceived materiality into the quantitative field of the machine. In this translation process, many essential characteristics of human listening are necessarily lost, but at the same time a new formal logic is acquired that is capable of operating on materiality in new ways. The alliance between human listening and experience and the quantifying and classifying capability of the machine thus becomes a key for finding new strategies for the organisation of electroacoustic sound.

This chapter will describe all the essential steps in the creation of a material database, motivating the design choices and identifying the fundamental characteristics, including qualities and biases inscribed in this type of creation; subsequent chapters will illustrate how material databases can be the starting point for implementing various music-making strategies, particularly in the field of algorithmic improvisation. The design logic of the material database can be ascribed to 'corpus-based' strategies, meaning those strategies that employ large collections of sounds and data (corpus) as a starting point to trigger musical processes. This research is strongly influenced by contemporary discourse on corpus-based music, in particular by the Fluid-Corpus Manipulation (FluCoMa) project, a project aimed at providing tools for corpus-based music and machine learning (Tremblay et al. 2022). The FluCoMa project made it possible to distribute on the main music programming software (Supercollider, Max/MSP, PureData) a library of extremely useful functions and tools for the creation, analysis, and management of sound databases, making the technical implementation of my material databases extremely easier and faster. From a technological point of view, then, this thesis is part of a broader set of contemporary practices, to which I hope to make a fertile contribution, not so much from a technical point of view – almost everything I will explain is easily implementable and does not in itself represent a technical innovation – but in the way theory and musical strategies are combined in my approach. I like to think of my music-making as a subset of corpus-based music, that is, a corpus-based approach to algorithmic improvisation where the corpus is a material database. The scheme illustrates the subsequent stages in the formation of a material database.



The first step in building a material database is to record the necessary sounds. First, it is necessary to identify a material identity and delimit its boundaries: for instance, do we want to have a database of water, or only of river water, or perhaps of gurgling water, regardless of its source? The choice in selecting and bounding a material identity is a subjective and creative operation that constitutes a fundamental aspect in defining what the result will be at the end of the database development. A database containing instances of a material identity that is too vague (e.g. 'any kind of plastic excited in any way') is likely to produce perceptual results in which the identity is not easily intelligible; conversely, a material identity that is too narrow (e.g. 'glass struck by a felt beater') will produce results that are too redundant. There must be a precise identity and at the same time sufficient acoustic variety in the sound sources. The most malleable material identities, those that tend to produce differentiated sound instances depending on the type of excitation they receive and the type of interaction they have with the space and the rest of matter, are those that most lend themselves to the creation of a flexible and expressive material database – for instance, let us consider the thousands of nuances that can be obtained by rubbing, scraping, tearing, and crumpling various types of paper of different weights and consistencies, the identity of paper always remains quite intelligible, while the potential dynamic and spectral expressions are very broad. According to the type of material identity chosen, distinct recording possibilities will arise. An environmental material identity, such as wind, will require field recordings to be made in multiple places and at multiple times, while a more manageable identity such as paper can be recorded in the studio. The material identities recorded in the studio can be isolated from the context in which one would normally hear them. This type of abstraction would seem to support an acousmatic Schaefferian perspective, in which sounds are isolated from their context in order to forget their source; on the contrary, in this case the studio recording is a way of capturing as clearly as possible only one specific semantic aspect of sound, namely its material identity. It is therefore not a way of isolating oneself from the material nature of sound, but of unnaturally emphasising it to make it more prominent. The deliberate choice of trying to expunge the space from the materiality is an evident contradiction of the notion of situatedness I discussed in the first chapter, but at the same time it is a form of compromise that is instrumental for approximating certain material identities in the digital domain. Whereas humans optimally identify and contextualise entities and identities in a situated environment, machines struggle to decompose complex signals into different semantic unities. The spatiality and the active presence of a surrounding context can easily obliterate material identities and make them hard to be picked up by a microphone or to be recognised by an algorithm, therefore artificially marginalising the context can be a way of focusing on materialities which would otherwise be ungraspable. From this point of view, very close miking techniques also allow the spatial components of the environment to be minimised in the recording. When the material identity I have decided to study can be recorded in the studio, I tend to prefer close miking, because a totally dry sound allows me to reconstruct a spatiality afterwards without the influence of the original space. Since the recorded sounds are going to be segmented and sorted in many ways afterwards, the temporality between distinct sound events is not important; there will not be a direct causal relationship between what happened first and what after. Sounds can be recorded multichannel, but mono recording usually provides more compact data. Whatever the number of channels in the material database, the creative operations that will be performed with it can have an output of any number of channels. Naturally, if one wants to preserve certain directional and spatial information, the most appropriate number of channels should be chosen. Personally, I prefer to work entirely in mono.

Once all the sounds have been recorded, it is essential to spend time listening to them carefully. Acousmatic listening permits to partially distance oneself from the history of the emission of those sounds and to ascertain what effect is perceived in a decontextualised listening session: are material identities still recognisable? Asking someone else what they perceive can be a good way to get a more objective viewpoint of what has been recorded: those who are personally involved in poietic processes lose the privileged point of view of those who have a disinterested experience of sound. Sometimes what the microphone picked up is a microscopic universe that is unknown to us before the recording, revealing new shades of materiality. While during the recording process an attempt was made to obtain the longest possible amount of recorded sound, it is then extremely relevant to decide what to select. Any redundancy will be a redundancy in the database, any shortage cannot be subsequently filled. All sounds will have to follow a certain recording standard in terms of dynamic range, signal-to-noise ratio, and stipulated minimum values of sampling frequency and bit depth. It may be desirable to remove all silences between sound events either manually or with an algorithm, making the database more compact and free of unnecessary data, while being careful not to remove lower-dynamic sound events, which are very important in order to maintain a broad and versatile expressive dynamic. The selection process is not only a technical process, but a truly creative stage in which the future of the material database is largely decided. Some sounds may be discarded for purely aesthetic reasons, while other sounds, perhaps accidental sounds, may be chosen for their incidentally interesting characteristics. It is also necessary to ask oneself what the chosen size of the database is. This choice should consider that most of the operations that will later be performed on material databases involve loading them into random access memory. As much as having an extremely large database can provide extensive creative possibilities, one should not underestimate the fact that the database should be a curated selection of material. My approach is to have databases that are as compact and expressive as possible, according to the idea of small data (as opposed to big data), an attitude that Rebecca Fiebrink (Vigliensoni et al. 2022) has accurately explored and to which I fully subscribe.

The curated selection of sounds that will form the database can be exported in a single audio file, which then has to be segmented into many fragments. Segmentation is a crucial stage in the formation of a material database, as it defines the time scale in which the database is expressed. I usually create several databases from the same corpus of records, each segmented with different criteria and parameters. A good method to break down the curated selection of sounds into individual fragments is to use auto-segmentation algorithms that employ machine listening. Since there is no fixed periodicity in the corpus, segmenting sounds within a certain periodic number of milliseconds is not meaningful. Transients and other acoustic variations must be measured in order to correctly identify distinct events. Sounds can be segmented by measuring their on-set, or by using a novelty feature algorithm. The latter measures the variation of a given descriptor over time (loudness, MFCC, spectrum, centroid, etc.) and segments the sound whenever the deviation of the descriptor's value exceeds a set threshold. Defining the threshold value is extremely relevant because it determines both the number of resulting segments and their average duration; the choice of a particular value is therefore as pragmatic as it is fundamentally creative. It is also possible to set minimum and maximum duration values for each segment, in order to avoid excessive temporal variation between the elements of the corpus. My material databases generally consist of a few tens of thousands of segments, each lasting between tens of milliseconds for more impulse-based and gestural sounds, and a few seconds for sounds of a textural nature. Having various versions of the same database segmented with different threshold values allows for control over multiple time scales, making it possible to choose between the dimension of musical phrases, that of individual events, or that of the micro-movements contained within each event. A sheet of paper being shredded can thus be understood as a single segment, or subdivided into dozens of fragments – when considered one at a time, they can become individual musical units, each having slightly different morphological characteristics from the others. A segmentation based on variations in amplitude, loudness, or on-set detection tends to establish the boundary between one segment and the next by following the transients of the sound, so it is suited for gestural sounds or sounds with strong dynamic discontinuity. On the contrary, segmenting sounds on the basis of spectral variations allows to obtain segments that have a certain internal activity in terms of amplitude, but at the same time maintain spectral coherence, thus being ideal for textures with inner movement. Static textures and purely flat sounds are not suitable for segmentation and are therefore difficult to integrate in the construction of a material database. Typically, a material identity recorded according to the described criteria features highly diverse sounds, ranging from short impulses to long continuous textures. The choice of descriptor used for segmentation with the novelty factor algorithm partly determines which morphological types will be most appropriately fragmented. It is not ideal to use different segmentation criteria for specific morphological subsets of the database, as the dissimilar results obtained from the various segmentations would risk being too heterogeneous and would therefore compromise the subsequent database building stages.

After segmenting the recorded sounds, it is time to analyse each individual fragment. The purpose of the analysis is to encapsulate as compactly as possible a collection of essential features of the sound from a quantitative point of view. All the classic descriptors employed in machine listening procedures can be used to perform the analysis: loudness, true peak, Fourier transform and related transforms, Mel bands, Mel-frequency spectral coefficients (MFCC), spectral descriptors (such as magnitude, flatness, centroid, spread, flux, skewness, kurtosis, slope, and roll-off), chroma, pitch tracking, sine decomposition, roughness, Bark coefficients, non-negative matrix factorisation, and so on. Each descriptor is suited to measure and represent a certain feature of sound. All descriptors inherently capture something related to materiality, as perceived materiality emerges from sound interdependently to all acoustic features. At the same time, no single descriptor is sufficient to approximate the material features of sound. Some descriptors, such as the Mel scale and MFCCs, take into account the physiological characteristics of human hearing, while other more sophisticated descriptors even try to approximate psycho-perceptive parameters (Kazazis 2020). It is evident that no descriptor could ever fully represent the subjective phenomenological experience of the individual. Nevertheless, it is worthwhile to identify the type of analysis that can best encapsulate information salient to the perception of the materiality of sound. Since no descriptor alone is sufficient to describe all relevant acoustic parameters, a good strategy is to use many descriptors in parallel, as suggested by James Bradbury (2021). In the context of this thesis, the most commonly used descriptors are the Mel-frequency cepstral coefficients. These are coefficients obtained by first analysing the spectrum with a Fourier transform, thus remapping the magnitude values obtained on the Mel scale, then applying a discrete cosine transform to the logarithmic values of the magnitudes. The values resulting from this operation are an extremely condensed and representative form of the sound's cepstral characteristics. Because of their compactness, MFCCs are among the most commonly used descriptors in machine learning, speech recognition, and music information retrieval tasks. Any number of coefficients can be calculated; the higher the number, the more accurate but less compressed the description will be, always taking into account the trade-off between spectral and temporal information that Fourier transforms imply. After some empirical experiments on the number of coefficients to be used, I chose a reference value of thirteen coefficients, as this proved to be a sufficient amount to describe the spectral profile of the sound without increasing too much the dimensionality of the analysis. Other descriptors that have certainly proved useful for material databases are the measurement of loudness and the spectral centroid. The remaining listed descriptors were employed whenever they were appropriate to the intrinsic nature of the database – for instance, databases containing material identities with discrete pitches require a more accurate analysis of the fundamental frequency and harmonic structure of sounds, so in addition to MFCC, loudness, and centroid, pitch-trackers and chroma measurements were also employed. Analysis can be performed at a given instant of the sound, but then it would not reflect the evolution of the sound over time; if the spectral density of a sound was calculated at the point of its transient, very distinct results would be obtained with respect to its decay. For this reason, it is important to measure the values of the descriptors several times over the duration of the sound at a given time interval, thus measuring the variation or the average over time. In the case of fragments lasting a few milliseconds, it is sufficient to average the values of the descriptors over time, resulting in a single, statistically representative value.

However, when sounds unfold intricately across longer durations, it may be relevant to analyse the trend of the values over time by measuring the standard deviation, the derivatives, or the symmetry of the distribution by measuring skewness. Gathering this set of values can significantly improve the sound representation obtained from the analysis, but at the same time increases the number of parameters and values for each individual sound analysed, greatly expanding the volume of data. In most cases, the material databases created for this thesis used only the average value of the time descriptors, in order to make the data obtained more manageable and concise.

Analysing all the individual segments with the chosen descriptors results in an array of data for each segment. Every segment of the database has an index value, and descriptor values are associated with that index. Each descriptor constitutes a dimension of the array. Just measuring MFCC, loudness, centroid, and pitch forms a fifteen-dimensional array, which can be thought of as a multidimensional space. Each sound is represented by a point positioned in the multidimensional space. In this way, the collection of segments has been placed in a vectorial parametric space, and is thus potentially ready to be used creatively according to quantitative or procedural logical criteria. It is evident that thinking creatively in a fifteen-dimensional space is a process that requires too much cognitive effort; moreover, such a high dimensionality makes any real-time processes that could be performed on the database more demanding in terms of computational energy. It is therefore necessary to find methods to compress the data collection more effectively by reducing the number of dimensions. Dimensionality reduction is a mathematically complex task that involves a loss of information in any case. The issue is to establish a methodology that allows the data to be condensed without losing highly significant information. It is certain that in the large multidimensional collection of data associated with sound segments, there are irrelevant data that could be removed. Some measured values on individual points prove to be statistically irrelevant or redundant when compared to a large collection of related arrays. It is then possible to find strategies to remove that data and remap the collection into a smaller, but equally representative space. It is evidently a question of establishing an acceptable trade-off between size reduction and potential loss of relevant data. The FluCoMa library proposes the implementation of three types of dimensionality reduction commonly employed in data science, each having specific advantages and disadvantages: principal component analysis (PCA), uniform manifold approximation and projection (UMAP), and a more general algorithm for multidimensional scaling (MDS). PCA measures the variance in the data collection and reorders the dimensions by identifying the main components of the dataset and minimising redundancies. Whereas PCA does not perform well with overly diverse datasets, UMAP is particularly effective in such cases, because it tries to create low-dimensional spaces while preserving their original structure (McInnes et al. 2018). MDS, on the other hand, attempts to keep the distances between elements in lower dimensionality space constant and has distinctive results depending on which type of distance measure calculation is performed. A useful way to gain insight into what happens during the dimensionality reduction process is to plot each individual sound as a point in a two-dimensional graph. Each type of dimensionality reduction will rearrange the points differently. By comparing a plotting of the original data collection with the results of the various dimensionality reductions, it is possible to get an idea of what kind of operations were performed. Figuring out what happened to the data collection by judging the plotted results, and evaluating if the trade-off in data compression is acceptable, is a skill that has to be trained by empirical practice. Personally, I do not have a favourite algorithm, but I try to observe what effects each technique produces depending on the material identity I am analysing and choose accordingly. The actual amount of resulting dimensions is a parameter that has to be set by trying multiple possible configurations and looking at the results, as each individual database might yield completely different outcomes. Because PCA and UMAP procedures both involve a form of training, when finalising the dimensionality reduction process it is important to save the trained model, so that if new points have to be added or compared to the database or if an inverse operation of dimensionality de-compression wants to be performed, it can be done accordingly to what the reduction procedure did to the data collection.

It is critical to remember that the vector space produced by dimensionality reduction no longer follows the logic according to which a dimension corresponds to relative values of a descriptor, but it becomes an even more abstract form of representation of the analysis data. The new resulting dimensionalities therefore represent attributes that can no longer be directly ascribed to the features of sound – or rather, to features that the human being has deduced and intentionally identified through a precise mathematical operation. Instead, we obtain a space that is coherent only for that particular collection of data. Thus, if one listens to all the sounds in the database sorted from the lowest to the highest value of a single given dimension, one may or may not find a precise acoustic coherence, because a direct relationship between individual dimensions and sonic features have been sacrificed for a better multidimensional representation. If there is an interest in preserving a clear connection between acoustic parameter and dimensionality, then dimensionality should not be reduced in any way. In the context of my research, it is unimportant to preserve this relationship; on the contrary, it is more fruitful to explore the potential of creating new multidimensional spaces that exhibit peculiar properties.

After dimensionality reduction, it is essential to normalise the values to a scale suitable for human interaction, generally in the range between 0 and 1. A linear normalisation finds the minimum and maximum values of each dimension, remaps them to the required range and rescales all intermediate values to the new range while keeping the proportions intact. Since typically in a collection of sound analysis data there are always a few points that diverge significantly from the others – the so-called outliers – the linear scaling of values produces a totally inhomogeneous parametric space, in which a few sounds are found at the vertices of the multidimensional space, while all the others are clustered around the centre. It is also possible that the sounds are mostly condensed in an area other than the centre, in which case it is preferable to standardise the data collection so that the average value is at 0.5. To avoid the spatial distribution being so unbalanced and affected by outliers, instead of normalising linearly it is often more appropriate to ignore the extremes of the multidimensional space altogether. A 'robust scaling' is therefore obtained not by using the minimum and maximum values of the distribution, but by choosing non-extreme percentile values. In order to perform a proper robust scaling, it is necessary to observe the plotting of the data distribution before normalisation and to estimate the optimal percentile values.

Once the normalisation is complete, the database is ready. Three distinct elements are then obtained, which together form the database: the audio file containing all the sounds, a document indicating the value in samples where the start and end points of each segment in the audio file are located, and the arrays of values relating to each segment. Both the audio file and the collections of data stored in json files can be loaded into the RAM at any time during a performance with very low computational expense and minimal loading time, making material databases excellent tools for improvisation. Material databases can therefore be understood as foundational devices for many real-time corpus-based musical procedures.

Chapter 7: Mosaics of Superimposed Materiality

To condense a material identity into a database affords the possibility of thinking about the materiality of sound from new angles. Just as the formation of an identity occurs in the human mind through a plurality of different instances, in the same way the material database gathers and organises a multitude. While it may be possible to work with the materiality of sound without resorting to databases, these confer the advantage of being able to work on a plurality of nuances of the same material identity. Through a corpus-based approach, one can think of sounds *en masse*. Such a viewpoint has the advantage of allowing highly sophisticated articulations of sound, providing the essential structural elements for a syntax of sound events. Not only does the database collects sounds en masse, but it also comprises within itself a collection of analysis data with which it is possible to discover or invent new orders and new logical relationships between the sounds it contains. This is essentially the same idea as that underlying concatenative synthesis techniques, in which a given number of sound segments are recomposed in a new order with the aim of creating complex micro-structures. According to Diemo Schwarz:

> Concatenative sound synthesis methods use a large database of source sounds, segmented into units, and a unit selection algorithm that finds the units that match best the sound or musical phrase to be synthesized, called the target. The selection is performed according to the descriptors of the units (Schwarz 2006, 3).

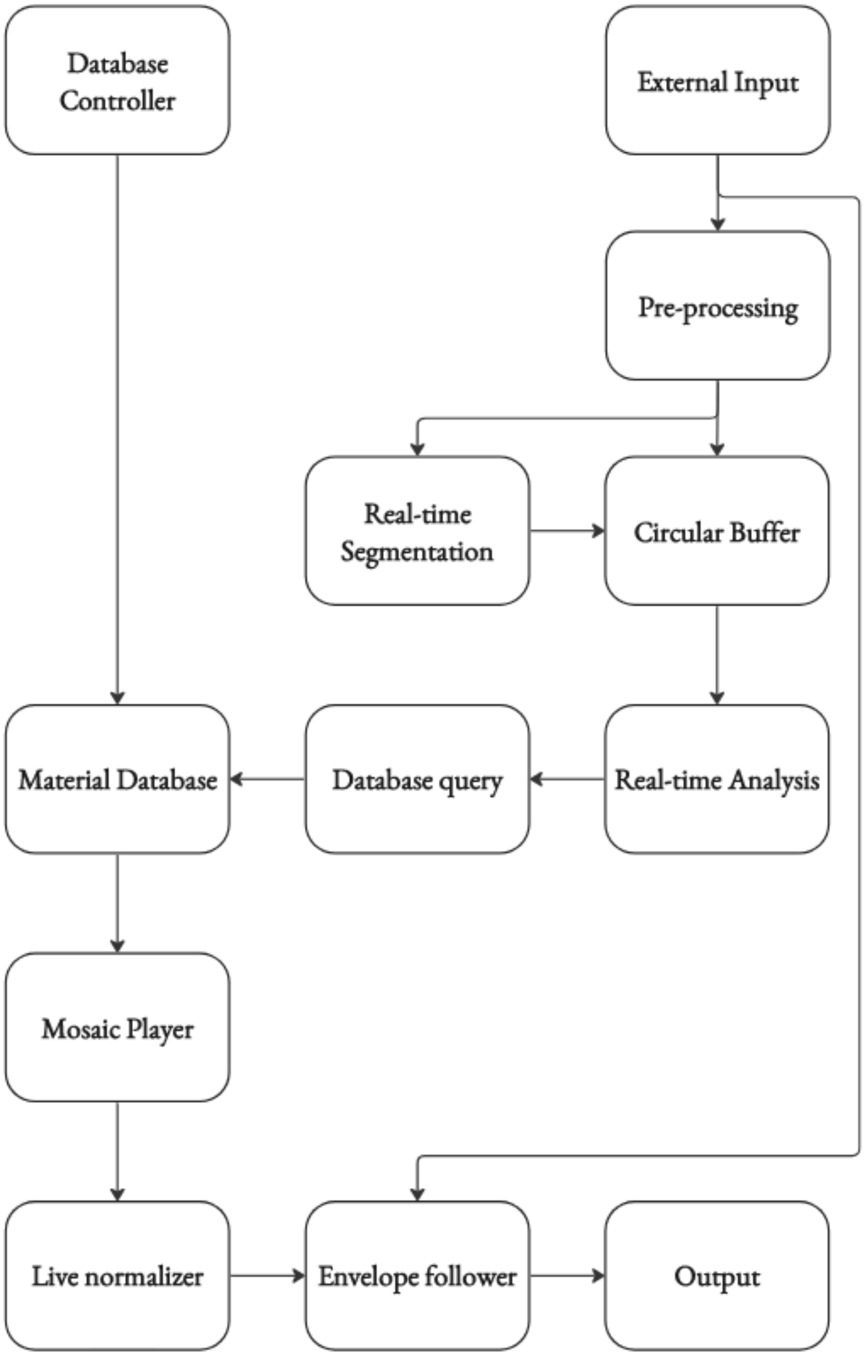
With a material database it is therefore possible to apply concatenative operations. The most relevant question is what criteria the unit selection algorithm uses to choose segments and organise them over time. Since the analysis needed to construct a material database requires projection into a multidimensional space, selection criteria can be established based on the vectoriality of this space. One can, for instance, trace trajectories in space and reproduce at an arbitrary time interval the sounds that correspond to the points in space traversed by the trajectory. I call this type of organisational procedure *sequential articulation* (listen to Appendix I for an example of this technique). The temporality of the sequence can be defined according to specific rhythmical choices, or thought in terms of varying densities. In this technique the criterion for selecting segments of the material database is abstract and is based on the idea of thinking of vector space as a space with certain intrinsic properties. Sequential articulation implies the belief that the vectorial logic of the database has properties that are meaningful and intelligible to the listener; or, if it does not, it requires a relinquishment to an arbitrary process in the hope that some logic will spontaneously emerge. While many algorithms for concatenating sound corpora are based on these principles, it is difficult to assume them in the context of research into the material identities of sound, because, as mentioned in the previous chapters, perceived materiality does not in itself have a vector continuity, nor can it have defined gradients. It is therefore inevitable that a geometric movement in the vector space of the corpus cannot correspond to an equivalent movement in the perceptual space of materiality. As much as it is possible to find coherent elements and potential correlations between a trajectory and its acoustic result, there is an underlying epistemological contradiction whereby there can be no univocal relationship between numerical representation and the phenomenological experience of listening. It is sufficient to test a sequential articulation algorithm on a material database to ascertain that there is not a consistent correspondence. Moving along geometric trajectories is therefore not a particularly meaningful process, although it can be used as a way of generating compositional materials or exploring the intrinsic characteristics of the database.

A more effective organisational strategy is to use the sound morphology itself:

> Dimensional complexity problematizes compositional control—one must devise a medium where morphology can be intuitively prescribed yet contain the level of detail required to represent the complexity of sound over time. […] an intuitive and exacting medium for prescribing morphology would be the use of sound itself (Hackbarth et al. 2013, 52).

Indeed, if the multidimensional organisation of the database was obtained by extracting data with sound analysis, it is probably the sound medium that is most appropriate to navigate within that space. By extracting information from sound sources and using it to control the selection of segments of a corpus, remarkably elaborate and meaningful sound articulations can be obtained. The application of this principle in the domain of material databases results in what I call *mosaic superimpositions* of perceived materiality.

Once a material database has been constructed, the implementation of the mosaic superimposition algorithm is relatively simple. This flow-chart shows a simplified scheme of the algorithm.



First, a sound stream must be identified to be used as a reference signal. Since this research focuses on improvisation and real-time techniques in general, it is appropriate that the signal comes either from the input of a microphone, or from a musical process that is taking place in parallel, or from another musician with whom one is playing. Eventually, the reference signal may come from an audio previously recorded and selected for this purpose. Regardless of where the signal comes from, it must be sent simultaneously to three different destinations: a real-time slicer or on-set detector, a circular buffer, and an envelope follower. The purpose of the slicer or on-set detector is to perform real-time segmentation of the signal, following criteria similar to those described for the segmentation of material databases. The threshold value, the type of analysis, and the time window of the analysis are parameters to be controlled in real-time, also with reference to the type of signal being received. The circular buffer takes care of recording the incoming signal. Whenever the slicer or on-set detector decrees the start of a new segment, the audio recorded in the circular buffer must be analysed immediately and the buffer must then be emptied to record the new input signal. The analysis performed on the content of the circular buffer must take place in the shortest possible time – usually extremely small time intervals, the minimum necessary for an FFT analysis window appropriate to the task. The descriptors used for the analysis must be exactly the same as those used for the creation of the material database, and be arranged in the same order, so that the array resulting from the analysis has the exact same structure as those contained in the database. If dimensionality reduction has been performed on the material database, then the array obtained by the analysis has to be reduced using the same trained model. Normalisation also needs to be performed in reference to the database. Each segment is thereby analysed, reduced, and normalised, and the resulting arrays are sent one at a time to an algorithm that compares them to those contained in the target corpus. The task is to compare the point described by the real-time segment array to the points contained in the database and to detect the closest one. A k-dimensional tree quickly identifies the nearest neighbour and pass its index value to an audio player. The segment in the material database that most closely resembles the segment analysed by the real-time signal is then played immediately, with a latency value equivalent to the duration of the analysed segment. It is possible to partially remove the original envelope segment from the reproduced sound with a live normalisation algorithm (Cipriani and Giri 2018) and then apply the envelope of the reference signal with an envelope follower. Applying this series of operations on each segment results in a concatenative stream of sounds belonging to the same material identity that mimic the behaviour of the reference sound. Since the resulting stream is made up of many small discrete sound events, I call this stream a mosaic, and since the mosaic forms a layer of materiality that can potentially be overlapped with the reference sound, it is a superimposition.

Depending on its relationship with the reference sound, the mosaic superimposition technique can produce various sonic outcomes and fulfil diverse musical functions. Appendix II-VI show sound examples from different applications of mosaic superimpositions. A first type of application occurs when the mosaic is used as an additional timbre overlay to expand the expressiveness of an amplified acoustic or electric instrument. The dynamic and spectral gestures of the instrument are closely followed by the mosaic. In this case, the mosaic superimposition represents a mere extension of the instrument, with which it entertains a symbiotic and strongly mimetic relationship. For instance, one can expand the timbral range of an electric guitar by applying a mosaic of an entirely dissimilar material identity to it, such as glass or water. Experiments of this kind were conducted within the scope of this thesis through rehearsals with various instrumentalists. The interesting aspect that emerged is that the instrumentalist's relationship to the mosaic is perceived as embodied, as there is always a direct causal correlation between instrumental gesture and sound response. In the case of amplified electroacoustic instruments like the electric guitar, it can sometimes be interesting to remove the original amplified sound of the instruments and use them as physical interfaces to play mosaics. Whether the sound of the instrument is amplified or not, mosaic superimposition allows for an extreme variety of material articulations of sound. It permits the player of an instrument whose material identity is static and clearly recognisable to be able to express themselves semantically by calling up different morphologies and identities. This type of semantic articulation of materiality is particularly effective when the reference instrument is the human voice, both because the human voice is extremely recognisable – and thus any superimposed layers are equally identifiable – and because listening to sounds of a non-vocal nature articulating typically vocal phonemes is particularly expressive. In addition, the voice can also use spoken language as a tool to create additional layers of meaning; a reciting voice can speak of drought while a mosaic of bubbling water is superimposed on its voice, creating an effect that is not only an extension of materiality, but of extra-coding. There are, however, instruments and instrumental types with which mosaic superimpositions combine less easily. Instruments that emit sounds of a textural, pitched nature, with shallow envelope shapes, such as the hurdy-gurdy, tanpura, and glass harmonica, are particularly difficult to segment. When an instrument has a static spectral and dynamic morphology, if the mosaic is trying to reconstruct the sound through a material database of a rather dissimilar nature, it will tend to almost always choose the same subset of segments from the database, making the resulting signal highly redundant. On the contrary, instruments with a wide dynamic expressiveness, especially with regard to the production of inharmonic sounds, are more suitable for superimpositions. The use of extended techniques on the instruments consequently also produces a wider range of responses from the mosaic.

A different paradigm of thought contemplates that the mosaic is not a mere addition to the instrumentalist, but rather a musical process relatively independent from it, with which the instrumentalist can relate in many ways. In this approach, the mosaic is the instrument that a performer plays and with which they interact with other performers. From a mere sound effect, the mosaic here becomes a proper algorithmic musical instrument that makes materiality and mimesis its fundamental expressive principles. The mosaic superimposition can be conceived as an essentially relational musical process by which a performer of algorithmic music can interact with any musician in the electroacoustic domain. Since the focus of interest in this thesis is on the expressive potential of perceived materiality in algorithmic improvisation, this approach is certainly favoured; it conceives the mosaic superimposition as a musical process interdependent on the other instruments, through which a performer of algorithmic music can express themselves and engage in dialogue with other musicians. The paradigm that sees the mosaic as a layer of the instrument confers to the mosaic an essentially passive role, as a mere addition to the timbral and material richness of the instrument. Differently, in this case the mosaic is actively played. There are several parameters that can be controlled during the performance. The threshold value of the segmentation of the input signal (i.e. of the other instrumentalists), the type of descriptor used for segmentation, and the time window of analysis influence the density of the mosaic and are therefore extremely expressive parameters. The duration of segments not only has to do with temporal organisation, but also influences which sounds in the database will be played. Similarly, processing the input signal with low-pass, high-pass, or band-pass filters is an effective technique to bias the sound before it is analysed, so that specific areas of the spectrum are emphasised. The real-time selection of different material databases with varying threshold values for segmentation, on the other hand, is the parameter that has the greatest influence on the material typologies produced. The possibility of alternating between the envelopes of the mosaic sounds and that of the incoming sound by means of an envelope follower is able to create various levels of dynamic similarity. Furthermore, the sound output of the mosaic can be transformed with all sound processing techniques usually employed in the electroacoustic domain. Simple filters applied to the output, when properly related to the filtering of the input sound, allow the creation of composite spectral movements, making the mosaic particularly expressive. Additionally, the use of a delay before analysis creates a time difference between the incoming sound and the mosaic imitation. One can think of this time delay as a way of producing call-and-response sound exchanges, but also as a way of producing canons of materiality. It is therefore possible to think of interaction with other musicians in terms of *material counterpoint*, a type of counterpoint in which imitative musical lines differ from each other not in terms of pitch, but in their perceived material characteristics. This perspective is also potentially interesting for compositional applications, where the mosaic can provide contrapuntal functions in the development of materials.

A third approach to mosaic superimpositions is to play them without interacting with other musicians. This configuration can be achieved either by using a dummy audio or by feeding the mosaic with sounds incoming from other processes. The use of audio dummies permits to produce imitations of a reference sound track, without it being made explicit. For instance, a speech track taken from a radio broadcast can be used as input to activate a mosaic of wood sounds, but without the speech being amplified. Using dummy sounds produces a morphological, spectral, and dynamic imprint that allows mosaics to be animated by borrowing their morphosyntax from the reference. The dummy sounds should therefore be chosen not for their semantic meaning – as this would remain outside the listeners' perceptual field anyway – but for their spectromorphological characteristics, as these will largely be transmitted to the mosaic. It is also possible to work in this way to produce extreme forms of remixing of musical compositions by other composers. Another type of application is audiovisual, in which audio from a video source is replaced by a mosaic imitating it with another materiality. I spent several hours watching videos of fireplaces where sound was replaced in real-time by mosaics of crumpling paper. The kind of perceptual illusions that can be achieved in audiovisuals using material substitutions through mosaics is quite broad and falls within the realm of post-acousmatic practices.

If, alternatively, mosaics are used in relation to other sound processes occurring in parallel, mosaics can be thought of as adaptive processes that dynamically react to the input. If the interaction between a mosaic and another sound process is reciprocal, then feedback and dynamic equilibria can be created. A mosaic containing aluminium foil sounds can imitate the sound of a filtered noise generator, whose cutoff and amplitude values in turn can be determined by analysing the sound output from the mosaic. Feedback can also be created between several mosaics, causing them to listen to each other; a dummy pulse of a few milliseconds is sufficient to activate the feedback chain. Naturally, a feedback between two or more mosaics after a while of chaotic fluctuation tends to get stuck on the reiterated playback of the same sound segments, but it is sufficient to alter any of the parameters of one of the mosaics to trigger an immediate transition to a new chaotic state. Due to the complex nature of these relationships, it is necessary to work carefully on the calibration between different mosaics before obtaining potentially significant results. These forms of mutual interaction can produce particularly suggestive sound articulations due to the semantic implications of the material databases used, such as the perception of a dialogue between two or more material identities, like rustling leaves and scraped paper imitating each other.

In conclusion, the mosaic superimposition algorithm lends itself to a plurality of modes of expression, from collective improvisation to cybernetic relationships between algorithmic processes. In each of its applications, mosaic superimposition enables the articulation of sound processes based on perceived materiality. Through mimesis, a constellation of potential relationships between material identities unfolds. Despite its expressive potential, the algorithm inherently has a number of obvious limitations that must be taken into account. A structural issue of mosaics is the way they relate to time, which is always discretised into segments. The partitioning of sound into discrete events operates in a perceptual domain that in some respects resembles that of granulation and presents all the morphological variety characteristic of concatenative techniques; at the same time this partitioning makes the signal discontinuous by definition. Mosaics are not at all accurate when it comes to imitating textural sounds, gradual pitch variations, and periodic acoustic phenomena, whereas they excel with fragmented or strongly turbulent signals. A certain continuity of the signal can be achieved by using segments of longer duration than the interval between one segment and another, by creating polyphonies of mosaic tiles, or by implementing interpolation systems between one segment and another. Another strategy could be to artificially prolong the duration of the output sound with resonant filters or stretching techniques if the input sound had a certain temporal continuity, analysing its development with additional specific descriptors. However, the implementation of such systems could prove to be excessively convoluted and computationally inefficient. It is perhaps more appropriate to ask whether it is possible to invent other techniques of material identity articulation with a corpus-based paradigm that might produce sounds with greater morphological continuity over time. It is precisely this question that instigated the research discussed in the next chapter.

Chapter 8: Neural Network Material Resynthesis

Concatenative techniques have the structural drawback of relying on discrete elements, with which it can be difficult to articulate the sound with a perception of temporal continuity. Although the fragmentary character of mosaics can certainly be a valid aesthetic choice in many situations, there are others where it would be more desirable to adopt instruments with greater continuity. As long as employed techniques are based on the reproduction of previously recorded samples, obtaining a sonic flow that is simultaneously flexible, dynamic, and temporally continuous is far from simple. In contrast, most sound synthesis algorithms have the advantage of relying on the production of continuous signals, so it is relevant to question whether they could be useful in the context of exploring continuous or textural material identities. Yet, precisely because conventional synthesis algorithms such as additive and subtractive synthesis are based on the periodicity of waveforms, they are generally poorly expressive in terms of perceived materiality. Some classical algorithms, such as FM synthesis, are capable of producing quite varied material types. Nevertheless, the parametric space they present turns out to be terribly inadequate for reasoning in terms of materiality: trying to produce a certain material identity with filter cutoff, frequency modulation ratio, or vibrato amount as controls is anything but straightforward. Even though it is not impossible to reach this goal, the task requires to constantly translate the thinking of perceived materiality into a parametric space which is designed with a completely different aim. This kind of mediation between symbolic representation and expected phenomenological result is cognitively intense and therefore might not fit the performative context.

There is, however, another class of algorithms, the physical models, which would apparently seem designed for this type of purpose. Physical modeling algorithms are based on the concept of reconstructing and abstracting the physical processes underlying the production of a given sound type, creating a virtual object model whose control parameters coincide with the parameters inferred by physics. This purpose is apparently analogous, but rather dissimilar from that of this thesis. If physical models aim to faithfully reproduce the physical behaviour of sounds, in contrast this research is focused on the perceived aspect of materiality. These two aspects of sound tend to diverge much more than one might think. A modal synthesis algorithm specifies exactly the values of density, elasticity, size of the virtual object, so that the physical properties of a given material can be simulated. Yet, if one hears in blind listening the model of an aluminium plate, then that of a uranium one, then of bronze, tin, concrete, glass, plastic, and so on, it is extremely difficult for the listeners' identification to align with what the model claims to reproduce. Working on the materiality of sound with physical models has the risk of underestimating the discrepancy between the physical representation of reality and its phenomenological perception. Since physical models aim to reproduce a defined physical entity, they must necessarily resort to a type of ontology in which it is always an object that is studied. Instead, perceived materiality is based on perceptual schemes in which there is not necessarily an object, but sometimes a phenomenon, or a field of relations between multiple entities and phenomena. In addition, physical modeling techniques have traditionally focused on reproducing the behaviour of musical instruments, a subject that is scarcely relevant within the scope of this thesis. It is only in recent years, thanks to the field of applied audio in film and video games, that more attention has begun to be paid to non-instrumental material processes. There are currently physical models of certain material objects that are highly convincing from a perceptive point of view. However, the implementation of each of these objectualities requires technical skills of a high engineering level, which are hardly available to a musician wishing to develop their own models. In my personal exploration of physical models, I observed that the most convincing models either consist of offline tools or require a great computational cost that can hardly be handled in a laptop performance; at the same time, the more viable models, such as waveguides and modal synthesis, either tend to produce redundant behaviours or exhibit strong non-linearities that could produce unpredictable responses and even break the instrument. Therefore, I believe that physical models present too many problems, both conceptual and technical, for them to be part of an algorithmic improvisation approach focused on perceived materiality.

If reconstructing the physics of sound is an overly cumbersome and complex task, a good strategy might be to infer material features directly from the analysis of acoustic sounds. If the properties of a synthetic sound could be easily modulated using those same sonic characteristics extracted from a reference sound, it would be possible to model continuous sounds with the most disparate material qualities. This type of task, pertaining to the field of descriptor-based synthesis, is far from simple. The most straightforward method is to collect highly descriptive analysis data of the reference sound, for instance by using several parallel FFT analysis bands, and apply the values obtained from the analysis to a multiband equaliser that models the amplitude of a noise in the various zones of the spectrum. This produces a sound that roughly follows the pattern of the reference signal, but has a rather different perceived materiality, entirely attributable to the noise generator that produced it. For a synthesised sound to have a convincing materiality borrowed from an acoustic sound, it would require a large number of analysis bands; however, due to the typical Fourier transform trade-off, such a large number of bands in the frequency domain would result in an unacceptable loss of detail in the time domain. In order for a synthetic sound obtained from filtered noise to be able to convincingly assume the material characteristics of a reference sound, it is necessary for the synthesis algorithm to be able to infer the morphological characteristics to be assumed from a smaller set of information. It should be able to reconstruct from a reduced number of data a larger set of values to assign to the synthesis parameters, so that the result is perceptually convincing. A synthesis algorithm can only perform inferences and regressions if it has a memory and is able to train with that memory. A resynthesis of perceived materiality therefore calls for the use of machine learning algorithms. Neural networks excel in learning certain numerical configurations, because their structure made up of layers whose bias values recalibrate is an effective pattern-matching tool. Having a sufficient number of meaningful data about a dataset of reference sounds, a machine learning algorithm is able to infer which analysis values are meaningful and how they should be related to each other; the obtained control values are applied to the synthesiser, the output of which is analysed to check the degree of similarity with the sounds in the dataset. Each individual attempt tunes the complex equation governing the relationship between the dimensional space of the analysis data and that of the synthesiser's control parameters. After millions of attempts, assuming the input data were sufficiently descriptive, a synthesis model capable of replicating the characteristics of the sounds contained in the database is obtained. Applying this type of technique is extremely significant in the context of research into perceived materiality, as the training dataset can be a material database. This allows to train synthesis models that mimic the characteristics of a given material identity. The models obtained from this process can then be used for many musical applications.

In recent years, numerous neural network synthesis algorithms have been implemented, including WaveNet (Oord et al. 2016), SampleRNN (Mehri et al. 2016), NSynth (Engel et al. 2017), MelGAN (Kumar et al. 2019), and DDSP (Engel et al. 2020). Each of these algorithms tried to find specific solutions to compensate for structural problems, such as the excessive amount of computation required for training, the effectiveness of data encoding and decoding, the sound quality of the result, and noise reduction. Given the enormous amount of data and calculations required for training, most of these implementations work at extremely low sampling rates and bit depths, resulting in sonic outcomes that are not particularly convincing in terms of perceived materiality. The computational complexity of these processes also makes them fundamentally impractical for those without high-performance and expensive GPUs. Recently, the possibility of renting a GPU via cloud computing systems has reduced the costs required for training. Given the collective interest in artificial intelligence in the field of digital arts over the past year, research has accelerated rapidly, making the models much more efficient and yielding more satisfying results. For the type of task required in this thesis, by far the most effective and accessible implementation is that proposed by the Realtime Audio Variational autoEncoder (RAVE) developed at IRCAM by Antoine Caillon (Caillon and Esling 2021). RAVE is able to train neural audio synthesis models up to 48kHz having perceived properties extremely similar to those of the sounds contained in the dataset. Once a model is trained, it is able to imitate an input signal using the information it has learned from the dataset. This type of operation, generally thought of for timbral transfer applications, becomes particularly interesting when conceived in terms of *material transfer*, in a similar way to the mimesis performed by mosaic superimpositions. RAVE's architecture is based on two learning stages, a representational phase and an adversarial fine-tuning phase. During the first phase, the sounds contained in the dataset undergo a multiband decomposition of the raw waveform by means of Pseudo Quadrature Mirror Filters (Nguyen 1994), so as to obtain an extremely condensed representation in only sixteen bands; the acquired values are then used for the training of a convolutional neural network with a 128-dimensional latent space, which trains a simple classical synthesis algorithm to behave like the sounds contained in the dataset. The synthesis algorithm combines a waveform generator, an envelope generator applied to it, and a multiband filtered noise that is added to the output to improve the representation of noisy sounds. The output of this first step is still rather rough and presents many imperfections, so the adversarial phase aims to improve the output through training with Generative Adversarial Networks

(GANs). After three million steps, the training is completed and the model is exported in TorchScript format and can later be loaded into Supercollider or Max/MSP for real-time use. The trained model can receive incoming sound that will be decomposed as in the training phase and based on the analysis data will produce a signal related to it. Since a dimensionality reduction operation has been applied to the model, in the phase between encoding and decoding of the real-time signal, it generally has a number of dimensions of less than ten. Each dimension can be altered manually by introducing a bias to the input signal. The quality of the output sound is strictly dependent on the type of dataset provided, the dimensions, and the amount of steps performed.

For the research relating to this thesis, I made several training attempts, which resulted in the creation of two models, one respectively trained to produce the sounds of gurgling stream water, the other to imitate the sounds of fireplaces (listen to Appendix VI-VII). In addition to evident symbolic and semantic reasons, these two material identities were chosen for their acoustic properties. Both crackling fire and gurgling water have a wide spectral and dynamic range, great internal activity, and at the same time considerable temporal continuity. Given their morphological plasticity, both of these two material identities are particularly suited to mimesis operations. Besides, there is no human being on the planet who does not have some notion of these two elements, so they are highly recognisable. The implicit aim of the training was also to produce two models that could interact with each other both morphologically and semiotically. The water model was trained first and required several months of testing before satisfactory results were obtained, while the second took a few weeks. It was important to establish a trade-off between the overall duration of the dataset and the training time required: the larger the size of the data, the more laborious the training, and also the more expensive, as it was necessary to rent GPUs through a cloud computing service. After several attempts at collecting reference sounds, I was able to condense the material identity of the gurgling water into a dataset of only half an hour, in which every single second was carefully selected, edited, and equalised. The selection criteria were based on representing a wide range of internal variations of the same material identity. It is important to emphasise that from the very beginning a very precise choice was made about the subjects to be represented: not water in general, but a specific type of acoustic behaviour produced by water in streams. Similarly, the preparation of the fire dataset focused on the collection of fifty distinctive records of fireplaces and small bonfires. The selection of behavioural subsets is necessary for a small dataset, because the more diverse the content of the database is in terms of sonic morphologies, the larger amount of data – and consequently a longer training time – will be required to obtain a convincing model.

Unlike the mosaics, in the training with RAVE the temporal articulation of the elements is relevant. Whereas for material databases the consequential order of events was indifferent, RAVE instead takes into account the continuity of events within a relatively short time window. When the analysis data are projected into a multidimensional space, the RAVE algorithm takes into account not only the data as points in space, but also the point-to-point trajectories between each consecutive sample of the analysis. It is precisely this feature that allows the model to have a coherent temporal organisation.

The results obtained at the end of the training are particularly surprising in the verisimilitude of the signal produced and the imitative ability of the algorithm. I performed blind listening tests with listeners, all of whom identified the sounds they heard with the material identities related to it, without recognising the use of digital tools. When informed of the artificial origin of the sounds, upon a second listening some of them reported experiencing a perceptual shift: knowing they were dealing with an artificial intelligence, they began listening with the expectation of recognising the algorithm's inherent material characteristics instead of the material identities it was trying to mimic. From a phenomenological-perceptual perspective, the conscious listening relationship with neural network material resynthesis can produce new (and potentially uncanny) identities that might expand the expressive field and open new questions and aesthetics. However one may speculate that such perceptions are influenced more by precise ideological expectations towards artificial intelligence than by phenomenological experience itself, it is in any case very fascinating to experiment with the potential perceptual ambiguity introduced by deep neural synthesis models. Trained models can imitate signals of other instrumentalists or any reference sound in real-time, although they have a latency time required for analysis in the order of hundreds of milliseconds. They therefore lend themselves to all the musical applications described in relation to mosaics, but produce distinctive perceptual results. Mosaics are relatively quick to implement, consume little CPU power, and produce sounds of a quasi-granular or otherwise collagistic nature; models trained with RAVE require a considerable investment of time and money to train, but produce temporally continuous sounds that accurately mimic the reference dataset. The synergy between these two different corpus-based material sound processing strategies allows to construct an articulate and flexible performative environment for algorithmic improvisation.

Chapter 9: Performative Strategies

The strategies described so far can be combined to achieve a more complex field of interactions and manipulations of perceived materiality. By bringing together mosaic superimpositions and neural network material resynthesis, a performative environment for algorithmic improvisation with multiple expressive possibilities can be developed. Instead of structuring a fixed environment, the implemented techniques are suitable for a bricolage programming approach, in which the system is reconfigured according to specific performance needs. The type of material databases employed and the ways in which the various material-based processes are related to each other changes depending on the type of situation and the possible relationship with other musicians. The techniques described in this thesis are regarded as building blocks that can be interconnected in a variety of ways, producing a multitude of results. What all combinations have in common is the fact that the parametric space I relate to is always oriented towards perceived materiality as a fundamental constructive element.

In the case of solo performances, the environment is configured to create feedback and complex relationships between different sound generators and imitators. A series of material identities are chosen for their semantic and morphological value, then material databases are put into mosaics and models are constructed with RAVE. Subsequently, the preparation of the performance focuses mainly on the type of sound interactions that can be obtained from the mutual imitation of the various processes. A mosaic of bubbling water sounds can interact with a neural resynthesis model also based on a water dataset, so that the same material identity is articulated in a contrapuntal manner by two different techniques – or two models, of fire and water respectively, can try to imitate each other by creating a dialogue between the two elements. Complex feedback chains can be structured in which several processes listen to and influence each other. A mosaic of paper sounds can produce a signal by imitating a bandpass-filtered noise generator, and the result can be sent as input to mosaic of aluminium sounds, which in turn will imitate the signal, and that will then be imitated by a mosaic of plastic sounds, whose centroid values control the bandpass frequency of the noise generator. In this type of composite chain of mimesis relationships, even a slight variation in the threshold values of an analysis can significantly affect the output of the system. Creating complex relationships produces a variety of unexpected results, so the analysis and processing chain must be carefully studied and balanced before performance. To play within concatenations of materiality, it is necessary to have a thorough understanding of the system's range of possible reactions. It is necessary to ascertain which configurations produce redundant attitudes and which ones produce perceptual chaos, in order to find the right balance between the two extremes. Once these relationships have been found, it is possible to conceive the control parameters as a multidimensional space that can be controlled by sensors or other physical interfaces. The most interesting type of sensor I have experimented with is the Leap Motion, a controller that tracks the behaviour of hands and individual fingers through recognition algorithms applied to infrared camera images. Leap Motion, as well as other similar interfaces, are particularly expressive from a performance point of view because they allow hand movement in space to be used as choreography to control sound. From a semantic point of view, the idea of performing tactile gestures in connection with the production of sounds endowed with particularly synesthetic material characteristics produces several levels of meaning simultaneously. In this situation, on one hand there is the semiotic relationship between the gesture and the supposed nature of the sound's original emission – for instance, the crumpling of paper, which is a gesture that typically reveals a tactile interaction – and on the other hand, the gesture is performed in the air, without any object, so there is the metaphor of attempting to grasp the intangible. If the materiality of the sound then changes and becomes something different, which does not directly recall a manual control, but at the same time it is the performer's hands that shape the sound, then a further level of meaning is created. By playing with this type of relationship between movement and the resulting sound, post-acousmatic music performances can be developed that explore the intermedial relations between sight and hearing through notions of touch.

When playing together with other musicians, the strategies of material identity articulation must be embedded in a specific performative framework, which takes more account of the relational aspect. In this context, mimesis is the main mode through which one can build a relationship with other musicians. Mosaics and neural network material resynthesis models can have the function of expanding the perceived materiality of other instruments, the latter being responsible for the gesturality of sound. A percussionist can play different types of objects while their perceived materiality is augmented by the superimposition of layers obtained from various material databases. One can play with the semantic aspect that is superimposed on the reference sound. In this type of situation, the role of the person handling the materiality articulation algorithms can be interpreted as purely passive, or merely responsive, but there is always a relevant margin of expression resulting from the control of the real-time sound analysis and the type of material processing employed. In my opinion, interaction with other instrumentalists is interesting when there is a gradient of possible relationships between their sounds and those produced by my mimesis processes. Strictly imitating the sound of others, creating counterpoints with temporal latencies, obtaining parallel sound processes from the combined use of various techniques, or even using material databases in a non-imitative manner, are all tools available for a performance. The aspect to pay attention to is mostly to understand what degree of independence or relative dependence on other instruments one wants to have at what moment, and what sonic meanings it can produce. What I find stimulating about this kind of approach is that it can fulfil diverse musical functions, from imitation to opposition to dialogue. When I play with other musicians, I assemble a system in which I have various types of processes – mosaics, neural networks, and other sound processing techniques – in parallel; I construct the system in such a way that each individual process can interfere with the others simply by activating a switch, so that internal relationships can be created between the elements. However, in this setting much of the focus is on the relationship with other performers and the type of analysis that is employed. When I play with other musicians, the majority of the controls I have in my parametric space have to do with the analysis performed on the sound received from the others: threshold values, descriptor type, equalisation and filtering of the input sound, segmentation type, and so on. The manipulation of these parameters cascades a series of combined effects on mosaics and material resynthesis. The most effective way of controlling the musical processes centred on perceived materiality that I have implemented in this thesis is to model the incoming sound and the way the processes listen to it. Relating in terms of perceived materiality with corpus-based tools requires a perspective that is always relational, always causally related to what happens through the agency of others. In my previous experiences of collective improvisation I had the impression of being an agent immersed in a larger system, in relation to which I always had to find a way to make myself present without overshadowing the others, trying to direct the musical dialogue in a coherent way. When I play with a system structured from material databases, I instead have the feeling of already being within a flow of distributed processes. This flow is a network of sound interweavings to which I simply have the role of shaping what comes out of my system in relation to what goes in.

The inherently relational character of the processes described here is especially suited to performative contexts in which relationships are reciprocal, these being situations in which the performative systems of individual musicians influence each other. Of course, every instrumentalist is influenced by what other instrumentalists do regardless of the instrument they play, but here I am referring specifically to instruments that employ machine listening to extract information from an audio signal and use that data to model the behaviour of their own musical processes. These are therefore adaptive tools and environments, or furthermore, since the exchange is multidirectional, *cross-adaptive*. At the Institute of Sonology, I had the privilege of working extensively on this performative approach, forming a cross-adaptive ensemble with other students and former students, including Giulia Francavilla, Elif Gülin Soğuksu, Francesco Corvi, and Andrejs Poikāns. In various trio configurations, we tested a wide range of cross-adaptive performance possibilities, both in the studio and in concert. In our experiments, each performer receives the audio signal from each of the other musicians and has complete freedom to analyse it as they wish and map those values to their musical processes. We have tried to leave as much freedom as possible to each performer in the type of techniques and performance modes to be employed, while always maintaining a focus on the cross-adaptive aspect of the ensemble. The complexity of interactions created when several musicians form a network of adaptive relationships with each other could easily scale exponentially and become incomprehensible. Much of the work was therefore aimed at experimenting with as many interactions as possible and finding balances between our systems, empirically assessing which mappings, which types of analysis, and which processes interacted best with each other (Appendix IX is a sound example of a cross-adaptive call-and-response situation). In my personal experience, the use of mimetic instruments such as mosaics and neural network material resynthesis has proven to be extremely effective in their ability to react to external inputs and produce musically meaningful responses, as they are already oriented towards a relational dynamic with other sounds. These instruments allow me to focus on the perceived materiality produced by the ensemble within shared and distributed musical concatenations.

Ultimately, the techniques implemented as part of this thesis have proven to be versatile enough to be applied in a variety of performance contexts. In all applications, they made it possible to think of sound and performative action in terms of perceived materiality, making material identities the genesis of musical development. Thinking in these terms leads to a particular perceptual mode during performance; becoming aware of the inherent potential of a corpus-based approach to material identities induces a type of poietic listening focused on perceived material features. It is not the same kind of esthesic listening that characterised the incipit of the research through the subjective phenomenological discovery of perceived materiality. Differently, a poietic listening to materiality is a perceptual state directed towards the expressive potential arising from the performative environment. An action-perception feedback loop is created (Vaggione 2001), in which it is perceived materiality that conveys a reciprocal interaction between the performer and the output of the algorithmic environment. In the poietic listening of perceived materiality, the epiphenomenological-qualitative, morphological, algorithmic-procedural, and semantic-metaphorical dimensions coexist simultaneously. The multiple levels of analysis and experience of materiality all converge to inform a multidimensional perspective on sound, through which new aesthetic directions can be explored.

Conclusions

Reconsidering the entire itinerary of this research, one observes a trajectory that crosses numerous disciplines in search of a synthesis between thought and musical action. First, an experiential phenomenon was identified from a subjective point of view, the materiality of sound; then a faculty of human perception, perceived materiality, was recognised in that phenomenon, and a cognitive model of its functioning was established on the basis of contemporary scientific knowledge. An attempt was thus made to confront this knowledge with a semiotic perspective in order to investigate its relational and communicative character. A thorough questioning of the epistemological consistency of the notions of perceived materiality and material identities laid the groundwork for defining a methodology for employing these notions within the computational domain. The methodology was developed in relation to the artistic inclinations and expressive preferences resulting from my previous experiences, with the explicit need of overcoming the structural limits of conflicting paradigms. In the technological ambit, a way of thinking has developed that is capable of translating perceived materiality into musical processes, making use of the most appropriate contemporary techniques; this type of thinking both fulfils the aims of this thesis and proposes a possible perspective on some of the most recent directions in the research of sonic technologies. The implementation of a series of musical strategies, namely mosaic superimpositions and neural network material resynthesis, not only proposed technical solutions to problems of an expressive nature, but also prompted the development of a flexible and collaborative performative environment. It can therefore be acknowledged that the trajectory taken by this research has been able to combine phenomenological perception, theoretical models, technological designs, the search for new methodologies, and artistic expression, establishing continuity and complementarity between them.

This is only one of many trajectories that can be pursued. Many aspects were left out either because they deviated from the path, or because there was insufficient knowledge available. There are also critical issues, contradictions, and limitations implicit in many of the choices made and paradigms employed, as well as many unresolved questions that need to be explored. From the point of view of a theory of perceived materiality, any paradigm shift in the phenomenological approach or in the epistemology of cognition would yield completely different results. Moreover, as already explained, the sciences of perception that support this thesis are based on models that will certainly have to be re-evaluated in the coming years, in light of a multimodal perspective that integrates the perception of all the senses; post-acousmatic performance approaches that take into account the new perceptual models could emerge in relation to this perspective. There is also a need for more research into the semiotic aspect of perceived materiality from an inter-subjective point of view, studying the semantic implications of material identities more closely and analysing the effects in the esthesis of the audience. On the technical front, numerous improvements and new implementations can be made to further the insights contained in this thesis. Especially in the area of machine learning, more tests need to be performed, more extensive models need to be trained, and above all, the perceptual, poietic, and ethical implications of using neural networks need to be investigated more thoroughly. Given the interest that machine learning is having in the artistic field, it is evident that there will soon be a serious paradigm shift in audio technology. The changes produced by new techniques will offer new affordances and may suggest numerous applications in the field of perceived materiality. It can be assumed that it will soon be possible to perform timbral transfer – and consequently mimesis of materiality – with extreme ease and flexibility. The technical implementations employed in this thesis will therefore soon become obsolete, but the thought that developed them will not – on the contrary, it will be even more relevant – since it is a thought that seeks to find a synthesis between technique, perception, and performative action. Finally, the research conducted up to this point can be expanded beyond the notion of mimesis, towards an idea of the metamorphosis of sound, of a new alchemy of sound matter. This suggestion is saved for future research.

Appendix

I. The sound example 'Appendix_I_sequential' shows an example of a sequential articulation of a material database of water. The temporal logic of the events is not related to the intrinsic nature of the database. The sequence queries the database moving along the multidimensional axes, but there is no recognisable gradient.

II. The sound file 'Appendix_II_paper_guitar' contains an excerpt of a studio session with Hugo Ariëns. Hugo's guitar is on the right channel, while a mosaic superimpositions imitates it in the left channel. The excerpt exhibits the mimetic capability of the mosaic superimposition algorithm.

III. The example 'Appendix_III_foley_guitar' shows another fragment from a studio session with Hugo Ariëns where his prepared guitar is imitated by a mosaic superimposition made from a material database of various small foley sounds.

IV. 'Appendix_IV_chaya_paper' is a mosaic superimposition of paper sounds over an interview of composer Chaya Czernowin.

V. 'Appendix_V_lachenmann_paper' is an example showing a mosaic superimposition of paper sounds over an excerpt of a solo cello piece by Helmut Lachenmann.

VI. The soundfile 'Appendix_VI_efflussi' is a recording from a live performance titled 'Efflussi' in collaboration with Francesco Corvi. A mosaic superimposition of water imitates synthetic sounds following their spectral and dynamic contours.

VII. 'Appendix_VII_gurgling_voice' shows a trained model of neural network material resynthesis mimicking my voice, while I explain some aspects of the development of the model.

VIII. The example 'Appendix_VIII_firewater' is made from a poliphony of neural network material resynthesisers trained on gurgling water sounds that try to replicate the behaviour of a fireplace.

IX. A call-and-response canon among performers in a cross-adaptive ensemble with Giulia Francavilla and Francesco Corvi is shown in 'Appendix_IX_canon'. A trained RAVE model of water triggered by enveloped pink noise generates textures, with an harmony obtained by a set of comb filters, while the other performers analyse the output and use it for influencing their sounds. The frequencies of the comb filter are extracted by measuring the loudest partials from the other performers.

References

Adkins, Monty, Richard Scott, and Pierre Alexandre Tremblay. 'Post-Acousmatic Practice: Re-Evaluating Schaeffer's Heritage'. *Organised Sound* 21, no. 2 (August 2016): 106–16. https://doi.org/10.1017/S1355771816000030.

Ancona, Riccardo. 'Live Coding: Costruzione Di Un Sistema per l'improvvisazione Algoritmica'. Conservatorio Licinio Refice di Frosinone, 2020.

Anderson, Elizabeth. 'Materials, Meaning and Metaphor: Unveiling Spatio-Temporal Pertinencies in Acousmatic Music'. PhD, City University of London, 2011.

Atkin, Albert. 'Peirce's Theory of Signs'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University, 2023.

Bayle, François. 'Image-of-Sound, or i-Sound: Metaphor/Metaform'. *Contemporary Music Review* 4, no. 1 (January 1989): 165–70. https://doi.org/10.1080/07494468900640261.

Blackburn, Manuella. 'Working with Brevity: Short Soundfiles in Electroacoustic Composition'. In *International Computer Music Conference Proceedings*, 2019.

Bonnet, François J. *The Order of Sounds. A Sonorous Archipelago*. Falmouth: Urbanomic, 2016.

Bradbury, James. 'Harnessing Content-Aware Programs for Computer-Aided Composition in a Studio-Based Workflow'. University of Huddersfield, 2021.

Bregman, Albert S. *Auditory Scene Analysis. The Perceptual Organization of Sound.* Cambridge, MA: MIT Press, 1990.

Buss, Emily, Joseph W. Hall III, and John H. Grose. 'Development of Auditory Coding as Reflected in Psychophysical Performance'. In *Human Auditory Development*. Springer Handbook of Auditory Research. New York: Springer, 2012.

Caillon, Antoine, and Philippe Esling. 'RAVE: A Variational Autoencoder for Fast and High-Quality Neural Audio Synthesis', 2021. https://doi.org/10.48550/arXiv.2111.05011.

Calabi, Clotilde. *Filosofia della Percezione*. Roma: Laterza, 2009.

Chion, Michel. *Sound – an Acoulogical Treatise*. Durham: Duke University Press, 2016.

Cipriani, Alessandro, and Maurizio Giri. *Musica Elettronica e Sound Design 2 – Teoria e Pratica Con Max 8*. Contemponet, 2018.

Clopton, Ben M. 'Development of Perception: Psychobiological Perspectives'. In *Audition, Somatic Perception, and the Chemical Senses*, Vol. 1. Behavioral Biology. New York: Academic Press, 1981.

Cocker, Emma. 'What Now, What next— *Kairotic* Coding and the Unfolding Future Seized'. *Digital Creativity* 29, no. 1 (2 January 2018): 82–95. https://doi.org/10.1080/14626268.2017.1419978.

Crombie, Alistair Cameron. *Medieval and Early Modern Philosophy*. Cambridge, MA: Harvard, 1959.

Deleuze, Gilles, and Félix Guattari. *What Is Philosophy?* New York: Columbia University Press, 1996.

Demany, Laurent, and Catherine Semal. 'The Role of Memory in Auditory Perception'. In *Auditory Perception of Sound Sources*. New York: Springer, 2008.

Döbereiner, Luc. 'Towards a Materialist Conception of Sound as Thing'. *Organised Sound* 24, no. 2 (2019): 217–23.

Dufort, Louis. *Matériaux Composés*. DVD-Audio. Vol. IMED 0893. Canada: empreintes DIGITALes, 2008.

Eco, Umberto. *A Theory of Semiotics*. Bloomington: Indiana University Press, 1976.

Eco, Umberto. *Trattato di Semiotica Generale*. Milan: La Nave di Teseo, 2016.

Emmerson, Simon. *The Language of Electroacoustic Music*. London: Palgrave Macmillan UK, 1986.

Engel, Jesse, Lamtharn Hantrakul, Chenjie Gu, and Adam Roberts. 'DDSP: Differentiable Digital Signal Processing', 2020. https://doi.org/10.48550/ARXIV.2001.04643.

Engel, Jesse, Cinjon Resnick, Adam Roberts, Sander Dieleman, Douglas Eck, Karen Simonyan, and Mohammad Norouzi. 'Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders', 2017. https://doi.org/10.48550/ARXIV.1704.01279.

Fiebrink, Rebecca, and Baptiste Caramiaux. 'The Machine Learning Algorithm as Creative Musical Tool', Vol. The Oxford Handbook of Algorithmic Music. Oxford: Oxford University Press, 2018.

Gaeta, Livio, and Silvia Luraghi, eds. *Introduzione Alla Linguistica Cognitiva*. Rome: Carocci, 2003.

Ganchrow, Raviv. Private conversation, 17 November 2021.

Grill, Thomas. 'Perceptually Informed Organization of Textural Sounds'. In *OFAI Research Seminar*. Austrian Research Institute for Artificial Intelligence, 2012.

Hackbarth, Benjamin, Norbert Schnell, Philippe Esling, and Diemo Schwarz. 'Composing Morphology: Concatenative Synthesis as an Intuitive Medium for Prescribing Sound in Time'. *Contemporary Music Review* 32, no. 1 (February 2013): 49–59. https://doi.org/10.1080/07494467.2013.774513.

Harrison, Jonty. *Évidence Matérielle*. CD. Vol. IMED0052. Canada: empreintes DIGITALes, 2000.

Husserl, Edmund. *Ideas. General Introduction to Pure Phenomenology*. Oxfordshire (UK): Routledge, 2012.

Jung, Carl Gustav. 'Psychological Commentary on "The Tibetan Book of the Dead"'. In *The Collected Works of C.G. Jung: Complete Digital Edition*. Princeton (NJ): Princeton University Press, 2014.

Kazazis, Savvas. 'Psychophysical Scaling of Timbre-Related Audio Descriptors'. PhD, McGill University, 2020.

Kosuth, Joseph. *[Wittgenstein]*. 1990. Glass sculpture.

Kumar, Kundan, Rithesh Kumar, Thibault de Boissiere, Lucas Gestin, Wei Zhen Teoh, Jose Sotelo, Alexandre de Brebisson, Yoshua Bengio, and Aaron Courville. 'MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis', 2019. https://doi.org/10.48550/ARXIV.1910.06711.

Leibold, Lori J. 'Development of Auditory Scene Analysis and Auditory Attention'. In *Human Auditory Development*. Springer Handbook of Auditory Research. New York: Springer, 2012.

Lufti, Robert A. 'Human Sound Source Identification'. In *Auditory Perception of Sound Sources*, 2008th ed. New York: Springer, 2008.

Lynn, Stephens G. 'Noumenal Qualia: C.S. Peirce on Our Epistemic Access to Feelings'. In *Transactions of the Charles S. Peirce Society*, 21:95–108, 1985.

Magnusson, Thor. 'Scoring with Code: Composing with Algorithmic Notation'. *Organised Sound* 19, no. 3 (December 2014): 268–75. https://doi.org/10.1017/S1355771814000259.

McInnes, Leland, John Healy, and James Melville. 'UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction', 2018. https://doi.org/10.48550/ARXIV.1802.03426.

Mehri, Soroush, Kundan Kumar, Ishaan Gulrajani, Rithesh Kumar, Shubham Jain, Jose Sotelo, Aaron Courville, and Yoshua Bengio. 'SampleRNN: An Unconditional End-to-End Neural Audio Generation Model', 2016. https://doi.org/10.48550/ARXIV.1612.07837.

Merleau-Ponty, Maurice. *Phenomenology of Perception*. Paris: Gallimard, 1945.

Molino, Jean, J. A. Underwood, and Craig Ayrey. 'Musical Fact and the Semiology of Music'. *Music Analysis* 9, no. 2 (July 1990): 105. https://doi.org/10.2307/854225.

Nattiez, Jean-Jacques. *Music and Discourse: Toward a Semiology of Music*. Princeton (NJ): Princeton University Press, 1990.

Nguyen, T.Q. 'Near-Perfect-Reconstruction Pseudo-QMF Banks'. *IEEE Transactions on Signal Processing* 42, no. 1 (January 1994): 65–76. https://doi.org/10.1109/78.258122.

Nyström, Erik. 'Morphology of the Amorphous: Spatial Texture, Motion and Words'. *Organised Sound* 22, no. 3 (2017): 336–44.

Nyström, Erik. 'Topology of Spatial Texture in the Acousmatic Medium'. PhD, City University of London, 2013.

O'Callaghan, Casey. *A Multisensory Philosophy of Perception*. Oxford: Oxford University Press, 2019.

O'Callaghan, James. 'Mimetic Instrumental Resynthesis'. *Organised Sound* 20, no. 2 (2015): 231–40.

Onorato, Giovanni, and Riccardo Ancona. 'Raise the Curtain! A Critical Perspective on the Idea of Post-Acousmatic'. Presented at the Rethinking the History of Technology-based Music, University of Huddersfield, 2022.

Oord, Aaron van den, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. 'WaveNet: A Generative Model for Raw Audio', 2016. https://doi.org/10.48550/ARXIV.1609.03499.

Peirce, Charles S. *Peirce on Signs: Writings on Semiotic*. Edited by James Hoopes. Chapel Hill: University of North Carolina Press, 1991.

Robinson, William. 'Epiphenomenalism'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2019. https://plato.stanford.edu/archives/sum2019/entries/epiphenomenalism/.

Schaeffer, Pierre. *Treatise on Musical Objects: An Essay across Disciplines*. University of California Press, 2019.

Schwarz, Diemo. 'Concatenative Sound Synthesis: The Early Years'. *Journal of New Music Research* 35, no. 1 (March 2006): 3–22. https://doi.org/10.1080/09298210600696857.

Smalley, Denis. 'Spectromorphology: Explaining Sound-Shapes'. *Organised Sound* 2, no. 2 (August 1997): 107–26. https://doi.org/10.1017/S1355771897009059.

Solomos, Maki. 'Schaeffer Phénoménologue'. In *Ouir, Entendre, Écouter, Comprendre Après Schaeffer*, Buchet/Chastel. Paris, 2000.

Stavropoulos, Nikos. *Micro-Lieux*. CD. Vol. IMED 19159. Canada: empreintes DIGITALes, 2019.

Stockhausen, Karlheinz. *Four Criteria of Electronic Music*. Video. Oxford Union, 1972.

Stoljar, Daniel. 'Physicalism'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2022. https://plato.stanford.edu/archives/sum2022/entries/physicalism/.

Sturm, Bob L. 'Adaptive Concatenative Sound Synthesis and Its Application to Micromontage Composition'. *Computer Music Journal* 30, no. 4 (2006): 46–66.

Tremblay, Pierre Alexandre, Green, Owen, Roma, Gerard, Bradbury, James, Moore, Ted, Hart, Jacob, and Harker, Alex. 'Fluid Corpus Manipulation Toolbox', 7 July 2022. https://doi.org/10.5281/ZENODO.6834643.

Tye, Michael. 'Qualia'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2021. https://plato.stanford.edu/archives/fall2021/entries/qualia/.

Vaggione, Horacio. 'Some Ontological Remarks about Music Composition Processes'. *Computer Music Journal* 25, no. 1 (March 2001): 54–61. https://doi.org/10.1162/014892601300126115.

Vigliensoni, Gabriel, Phoenix Perry, and Rebecca Fiebrink. 'A Small-Data Mindset for Generative AI Creative Work'. Proceedings of the Generative AI and HCI Workshop presented at the Conference on Human Factors in Computing Systems Workshop, 2022.

Watsuji, Tetsuro. *A Climate. A Philosophical Study*. Japan: Ministry of Education, 1961.

Werner, Lynne A. 'Overview and Issues in Human Auditory Development'. In *Human Auditory Development*. Springer Handbook of Auditory Research. New York: Springer, 2012.

Wishart, Trevor. 'From Architecture to Chemistry'. *Interface* 22, no. 3 (1993): 301–15.

Wishart, Trevor. *On Sonic Art*. Amsterdam: Harwood Academic Publishers, 1996.

Yost, William A. 'Perceiving Sound Sources'. In *Auditory Perception of Sound Sources*. New York: Springer, 2008.

Page intentionally left blank